ЈУЛИ-СЕПТЕМБАР, 2022.

ИНФОРМАТИЧКА ТЕОРИЈА II

НАСТАВАК ИНФОРМАТИЧКИХ ПРИЛОГА

РАСТКО ВУКОВИЋ
НЕКАД. ПРОФ. МАТЕМ.
ГИМНАЗИЈА БАЊА ЛУКА

Растко Вуковић, јули-септембар. 2022.
Информатичка Теорија II – наставак информатичких прилога
https://www.academia.edu/86097551/

Предговор

Ова скрипта је непосредни наставак „Информатичке Теорије“[1], која је садржавала наслове, почев од „1. Неквадратна матрица канала“, до „43. Условне вероватноће“. Нумерација тема се наставља редом. Обе скрипте рађене су на основу сећања и записа питања и одговора на Блогу[2] актуелном у вези са мојом теоријом информације. Ти се прорачуни, објашњења и докази лако заборављају и потребно их је негде остављати „за касније“, верујем.

Текстови те прве скрипте (Информатичка Теорија I) завршени су и не намеравам их мењати какви год били разлози. Међутим, скрипта коју пишем, у текућим поглављима, увек је препуна лапсуса, рачунских грешки, погрешних претпоставки, на које се стално враћам и стално их исправљам. Као да више волим тражити и исправљати своје грешке него настављати са писањем, а заправо ради се о хипотезама које прво интуитивно постављам и током тестирам, дорађујем, или одбацујем.

Наводим то јер радне текстове држим на интернету, када ми понеки пријатељ шеретски добаци „Аха, нашао сам ти грешку математичару!“, а ја му узвратим „Добро, у текућој је верзији, али чик ми нађи грешку у завршеној!“ Овде сам заглавио са недоследним заокруживањем вероватноћа прво на две, па три, шест и више децимала и, наравно, са различитим резулатима (нарочито од поглавља регресија фреквенција слова), а онда због погрешних претпоставки и лошим оценама.

Аутор, 2022.

[1] Информатичка Теорија I, https://archive.org/details/it_20220709/
[2] Блог, http://rvukovic.net/blog/

Садржај

IV Део: Амплитуде

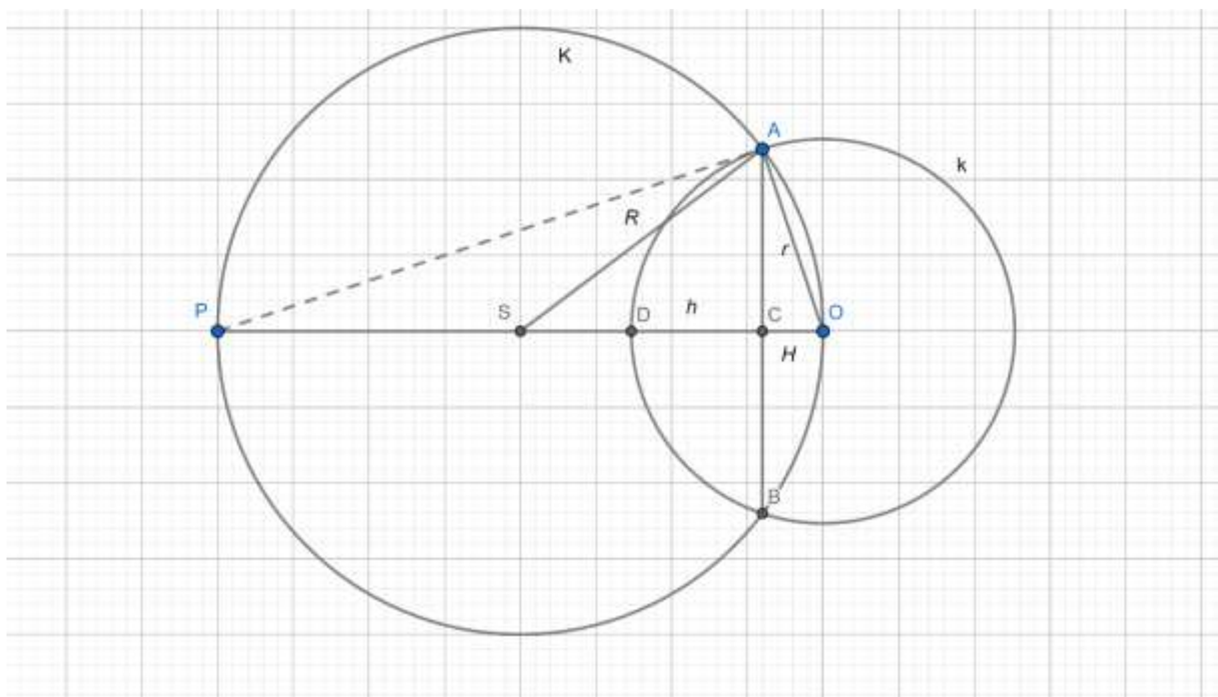
Две вредности карактеристичне за осцилације су таласна дужина и амплитуда. Прва повезује број, фреквенцију понављања основне фазе у јединици времена и пређени пут, дакле брзину таласа, а друга нам каже нешто о одступању фазе таласа од њене средње вредности. У оба случаја ради се о неким вероватноћама. Прве о неодређености налажења таласа на датом месту, а друге о снази израза таласа, његовој моћи да комуницира са околином, односно шансама опажања осцилације, или онога што талас представља од стране других.

Сматрам да Комптонов ефекат, поред познате интерпретације, говори о вероватноћи кретања са становишта саме честице-таласа (фотона), о неспонтаности скретања на мање густине положаја и праћењу таквих трајекторија као мање вероватних. Оно што се дешава током судара је деловање силе која (према мојој теорији информације) мења перцепције и вероватноће, након чега следи скретање и расипање просто као ново „највероватније кретање“ (са становишта учесника).

Мали је корак за физику до поменутог „информатичког“ тумачења Комптоновог расипања, а још је мањи до одговарајућег вероватносног тумачења таласне дужине и амплитуде (Wavelength II). Ако, када та становишта једном усвојимо, онда ће третирања тих таласних карактеристика добијати на ширини и, верујем, на начине који су овде тек започети.

44. Сфере потенцијала

Замишљамо две лопте, већу са центром S полупречника R и са центром O на њеној сфери мању лопту полупречника r. На слици је равни пресек тих лопти, по главној кружници K веће сфере и главној кружници k мање сфере. Пречник $P-S-O$ веће кружнице (K) је хипотенуза правоуглог троугла POA. Са периферије, тачке $A = K \cap k$, пречник PO види се под правим углом $\sphericalangle(PAO) = \frac{\pi}{2}$, а оштри углови, $\varphi = \sphericalangle(APO)$ у тачки P и $\varphi_1 = \sphericalangle(POA)$ у тачки O тог троугла, комплементни су, тј. $\varphi + \varphi_1 = \frac{\pi}{2}$. Централни угао веће кружнице $\sphericalangle(BSA) = \theta = \frac{\varphi}{2}$, где такође тачка $B \in K \cap k$. За тачке $C = AB \cap PO$, $D = k \cap PO$ требаће нам дужине $H = \overline{CO}$ и $h = \overline{DC}$.



Овако замишљамо простор вероватноћа, са груписањем вероватнијих догађаја све ближе центру веће лопте (S), при чему је њена сфера ма којег полупречника (R) увек граница вероватноћа. Она уравнотежује напетости унутрашње густине вероватноће и вањских сила неизвесности. Лопта са центром (O) на ободу веће представља (материјалну) тачку кратког домета деловања (r). Основне су теме (поводом претходних наслова) биле о релевантности односа дужина унутрашњег $\widehat{ADB}$ и вањског лука, посматрано из центра (O) мање кружнице (k), односно одговарајућих површина, те запремина мање сфере унутар и изван веће. У том тексту сам се определио за запремине, радећи по интуицији, али има смисла израчунавати односе површина, или дужина.

Пример 1. Дужина лука $(\widehat{ADB})$ кружнице (k) полупречника r је $l_1 = 2r\varphi_1 = r\pi - 2r\varphi = r(\pi - \theta)$. Остатак њеног обима је лук дужине $l_2 = 2r\pi - l_1 = r(\pi + \theta)$. Однос ових дужина, коефицијент, или „потенцијал“ $f' = l_1 : l_2 = (\pi - \theta) : (\pi + \theta)$, можемо апроксимирати, јер за наше потребе већи полупречник (R) много је већи од мањег (r). Из $\frac{r}{R} \to 0$ следи $\frac{1}{2}\frac{r}{R} = \sin\varphi \approx \varphi$, а отуда $\theta = 2\varphi = \frac{r}{R}$, па је $f' \approx 1 - \frac{2r}{\pi R}$. ❑

Пример 2. Израчунавамо однос површина делова мањег круга, заједничке унутрашњости $K \cap k$ према остатку. Површина одсечка (AOB) већег круга је $P' = \frac{1}{2}R^2 \cdot 2\theta - \frac{1}{2}R^2 \sin(2\theta)$, а површина одсечка (ADB) мањег је $P'' = \frac{1}{2}r^2 \cdot 2\varphi_1 - \frac{1}{2}r^2 \sin(2\varphi_1)$. То су разлике површине кружног исечка и троугла унутар. Овде је и даље $\varphi_1 = \frac{\pi}{2} - \varphi = \frac{\pi-\theta}{2}$ и $\theta = \frac{r}{2R} \to 0$, па налазимо:

$$P' = \frac{R^2}{2}(2\theta - \sin 2\theta), \quad P'' = \frac{r^2}{2}(2\varphi_1 - \sin 2\varphi_1)$$

$$P_1 = P' + P'' = \frac{R^2}{2}\left[(2\theta - \sin 2\theta) + \frac{r^2}{R^2}(\pi - \theta - \sin\theta)\right]$$

јер је $\sin x \to x$, када $x \to 0$, а $2\theta = \frac{r}{R}$. Тако ова P_1, заједничка површина кругова, постаје:

$$P_1 \approx \frac{R^2}{2}\left[0 + \frac{r^2}{R^2}(\pi - 2\theta)\right] = \frac{r^2}{2}(\pi - 2\theta) = \frac{r^2\pi}{2}\left(1 - \frac{1}{\pi}\frac{r}{R}\right)$$

Површина P_2 остатка мањег круга је:

$$P_2 = r^2\pi - P_1 = r^2\pi - \frac{r^2\pi}{2} + \frac{r^3}{2R} \approx \frac{r^2\pi}{2}\left(1 + \frac{1}{\pi}\frac{r}{R}\right)$$

Њихов количник

$$f' = \frac{P_1}{P_2} = 1 - \frac{2}{\pi}\frac{r}{R}$$

је коефицијент који дефинише „потенцијал“ у случају површина делова круга. Приметимо да је исте ознаке, јер је исте вредности као и количник дужина, онај из прошлог примера. ❑

Пример 3. Потражимо однос површина калота мање кугле, дела унутар веће и остатка. Висина калоте, капице која је део мање сфере у већој лопти:

$$h = \overline{DC} = r - \overline{CO} = r - r\cos\varphi_1 = r\left[1 - \cos\left(\frac{\pi}{2} - \frac{\theta}{2}\right)\right] = r\left(1 - \sin\frac{\theta}{2}\right) \approx r\left(1 - \frac{\theta}{2}\right)$$

па је површина калоте

$$P_1 = 2r\pi h \approx 2r^2\pi\left(1 - \frac{\theta}{2}\right)$$

Површина остатка мање сфере је:

$$P_2 = 4r^2\pi - P_1 \approx 2r^2\pi\left(1 + \frac{\theta}{2}\right)$$

па је однос тих површина:

$$f'' = \frac{P_1}{P_2} = \frac{2r^2\pi\left(1-\frac{\theta}{2}\right)}{2r^2\pi\left(1+\frac{\theta}{2}\right)} \approx \left(1-\frac{\theta}{2}\right)^2 \approx 1-\theta$$

$$f'' \approx 1 - \frac{1}{2}\frac{r}{R}$$

То је коефицијент који дефинише „потенцијал“ у случају површина делова сфере. ❑

Разлика коефицијената односа површине калоте дужина лукова је:

$$f'' - f' = \left(1-\frac{1}{2}\frac{r}{R}\right) - \left(1-\frac{2}{\pi}\frac{r}{R}\right) = \frac{4-\pi}{2\pi}\frac{r}{R} > 0$$

па је $f'' > f'$. У „Информатичкој Теорији“, у наслову „38. Виртуелне сфере“ за коефицијент, однос запремина заједничког дела и остатка мање лопте, који дефинише „потенцијал“ добијено је

$$f = 1 - \frac{3}{4}\frac{r}{R}$$

Поредак је $0 < f < f' < f'' < 1$. Свеједно, сва три коефицијента дефинишу у бити исту врсту потенцијала, они дефинишу силе које опадју са квадратом удаљености R. Најмањи од ова три потенцијала упоређен са релативистичким гравитационим даје:

$$f = g_{00}$$

$$1 - \frac{3r}{4R} = 1 - \frac{R_s}{R}$$

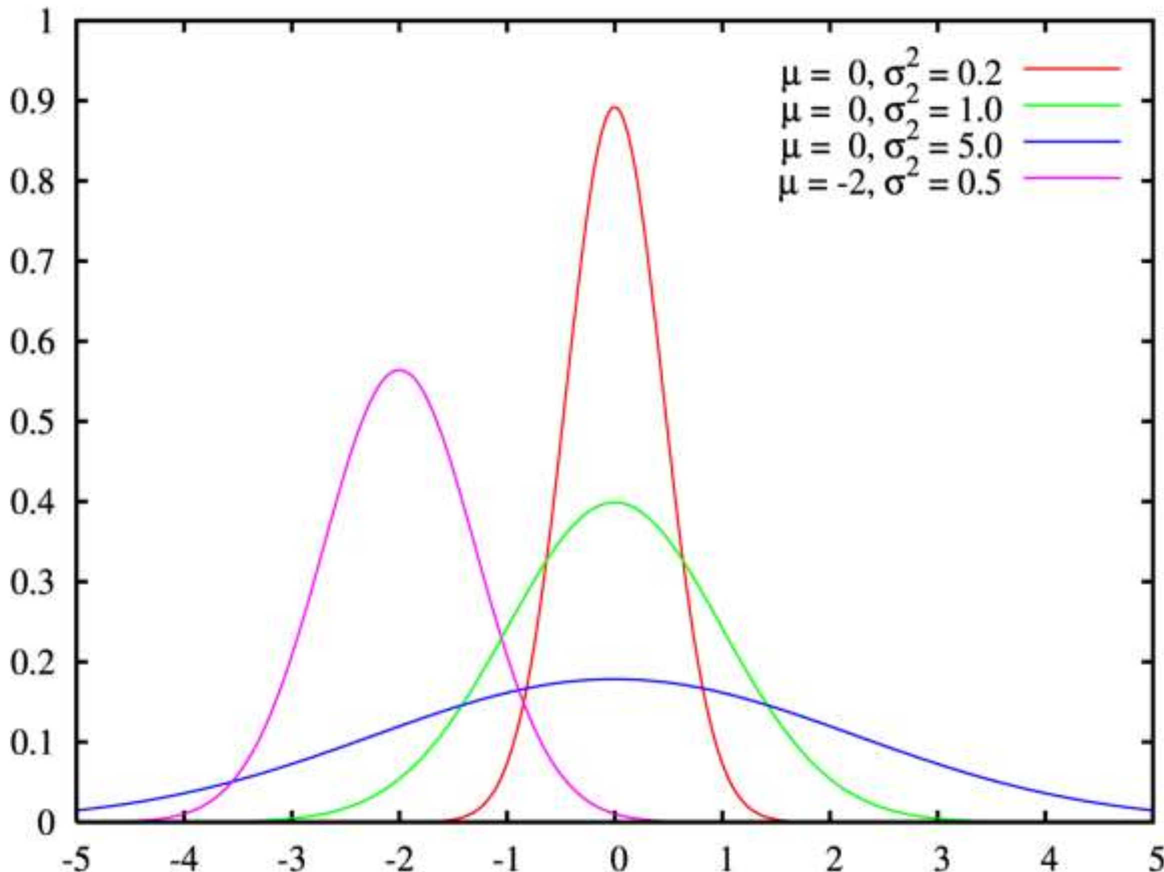
$$\frac{3r}{4} = R_s$$

где је $R_s = 2GM/c^2$ Шварцшилдов полупречник[3], у којем је G гравитациона константа, M је маса тела које гравитационо привлачи, а c је брзина светлости у вакууму.

[3] Простор-Време, 1.2.10 Шварцшилдово решење, http://rvukovic.net/knjige/Prostor-Vreme.pdf

45. Силе вероватноће

У апстрактном 3-Д простору вероватноћа замишљамо центре око којих се групишу извесности, постепено све већих вероватноћа што је удаљеност од центра мања. Та се идеја може разумевати помоћу графова расподела, рецимо Бернулијеве и њених апроксимација Пуасонове и Гаусове, који нам могу изгледати као да нека „сила“ нагомилава већи број исхода око средњих вредности.



На википедији можете погледати објашњење ових графова нормалне расподеле (Гаусове), ради инспирације, али дубљи разлог потраге за овим силама је у литератури подразумеван и никада до краја експлицитно изречен „принцип вероватноће“, да се вероватнији исходи чешће догађају.

Знамо да је и у физици упитан онај стари концепт силе из прошлих векова, који је за гравитацију Ајнштајн (1916) допунио закривљеним простором, дакле врстом „спонтаности“ које управо овде разоткривамо. Спонтано, чешће се догађају вероватнији исходи, топлота прелази са тела ниже на суседно тело више температуре, системи се настоје развијати ка мање информативнима, физичка кретања се подвргавају принципу најмањег дејства.

Сила о којој је овде реч припада поопштеној спонтаности, али опет, њене форме[4] проналазимо у Кеплеровом другом закону, у „кретању набоја“ по коникама, по деловању „брзином светлости“ и по опадању са квадратом удаљености. Сваки од тих феномена, аналогија, потребно је посебно и пажљиво преиспитати. Нешто од тога урадићу одмах.

[4] Uncertainty Force, https://www.academia.edu/83009873/Uncertainty_Force

Сфере потенцијала посматрамо ванвременски, као неку поставку позорнице на којој се тек требају појављивати неки догађати. Претпостављамо да је могуће такво безвременско, апстрактно стање и да је оно у том смислу врста математичке структуре. Међутим, црпимо инспирацију из физике.

У мојој теорији вероватноће, већ одавно рутински посматрам светлост, електромагнетне таласе и уопште фотоне, или друге објекте који би се кретали том брзином као безвременске, таквима који своју реалност дугују својим посматрачима. Објекти којима време не тече постоје у простору тамо су где им је највероватније место и биће на таквом месту и у следећем тренутку сваког релативног посматрача, у датој ситуацији коју може мењати само сила, а која је увек промена настала меном ситуација.

Сила мења вероватноће и промена вероватноћа нека су сила, становиште је које заступам својом теоријом информације. Поремећај вероватноће значи присуство силе – реченица је мојих текстова одавно, као и – чаша је на столу зато што јој је тамо највероватније место и биће тамо и убудуће све док на чашу не делује нека сила или неко друго тело, попут руке која ће чашу склонити.

У складу са реченим, безвременска тела тамо су где су, а посматрачи су ти који им „додају“ своје време, процесе и промене. Зато је брзина светлости независна од брзине извора, јер тек брзина промена онога ко светлост посматра даје њој брзину. Према томе, садашњост сваког од нас, сваке реалности, помера се дуж временске осе брзином светлости.

Аналогно, сфере потенцијала у свом математичком концепту статичне су, а динамику добијају тек опажањем од стране субјеката којима време тече. Обзиром на ток садашњости, од прошлости ка будућности, од исказане информације ка неисказаној, од реализованог ка нереализованом, сфере потенцијала биће виђане у ширењу.

Сви делови сваке поједине од тих „ширећих сфера“ истовремени су, што значи квантно спрегнути, те им у физичком смислу таласне дужине могу бити исте, а амплитуде све мање, као и другим тако ширећим таласима. То, пак, значи да свака од сфера носи исту „енергију“ или „импус“, али њене шансе интеракције са евентуалним другим објектом опадају са повећањем полупречника. Ово је прилично у складу са претходним, оним какве смо те сфере потенцијала откривали упоредо кроз праве физичке константне централне силе.

То нас доводи до „закона одржања“ вероватноће, такође подразумеваног у теорији информације (мојој), али раритета у званичној науци. Како онда такав повезати са нечим физичким – питаћете се, а одговор је – помоћу енергије. Како важи закон одржања енергије, или масе ($E = mc^2$), треба пратити израчунавања о вертикалном паду[5] као аналогију и добити нешто попут тамошњег

$$m = m_0 e^{GM/Rc^2}$$

где би маса m тела које пада била m_0 у одсуству гравитације. Развојем израза у Маклоренов ред и занемаривањем виших степена малих чланова, видимо да можемо писати

$$m = \frac{m_0}{1 - \frac{GM}{Rc^2}}$$

[5] Простор-Време, 1.2.8 Вертикалан пад, http://rvukovic.net/knjige/Prostor-Vreme.pdf

што значи да коефицијент повећања масе у гравитационом пољу

$$\frac{m_0}{m} = 1 - \frac{GM}{Rc^2}$$

макар формално одговара горе поменутом [44.] односу запремина f заједничог дела и остатка мање лопте, који дефинише „потенцијал вероватноћа“. Израчуната брзина сателита (у књизи из линка), је $v = \sqrt{GM/R}$, што долази из закона одржања енергије (масе) и преласка потенцијалне енергије гравитационог поља у кинетичку, приликом слободног пада.

Речником енергије, релативне E и сопствене E_0, долазимо до истог. Штавише, постоји аналогија између ових коефицијената опште и специјалне теорије релативности:

$$mc^2 = E = \frac{E_0}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}} = m_0 c^2 \gamma \approx m_0 c^2 \left(1 + \frac{1}{2}\frac{v^2}{c^2}\right) = E_0 + \frac{1}{2} m_0 v^2$$

Што је разлагање укупне релативне енергије ($E = E_p + E_k$) на сопствену, коју класична физика назива потенцијалном ($E_0 = E_p$), и кинетичку ($E_k = \frac{1}{2} m_0 v^2$). То што је данас и више од аналогије за теоријску физику, широмо даље на нови терен у теорију вероватноће.

Другим речима, за наводне потенцијале сфера вероватноћа можемо писати:

$$f \approx \frac{E_0}{E} = e^{-E_k/E_p}$$

а то значи да ће брзина v, која дефинише кинетичку енергију тела у слободном паду, одговарати случајној варијабли x нормалне расподеле вероватноћа. Коефицијент, или потенцијал сфере тако постаје

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

што је заправо густина нормалне (Гаусове) расподеле.

46. Информација сфера

Сфере потенцијала, којима се у овим узастопним насловима бавимо, статичне су. Попут таблице множења, оне су апстрактне конструкције без рецимо мењања током времена. Развојност њима долази од релативних посматрача чија времена „теку брзином светлости". Време (t) тако постаје имагинарна ($i^2 = -1$) дужина ($x_0 = ict$) која би се могла прећи брзином светлости (c). Димензија времена има више.

Свет сваког од релативних посматрача потенцијалних сфера структуриран је информацијом чија суштина је неизвесност, ради чега није могуће реализовати реализовану информацију. Поновљена „вест" није више вест, нити је са становишта истог посматрача узастопно виђење тачно исте сфере могуће. Сваки од посматрача стално се мења, односно његова околина, па онда ни он у њој никад нису исти. Сфере потенцијала које они могу видети шире се, титрају, носе потенцијал информације и понекад интерагују.

Укупна површина сфере расте са квадратом полупречника ($\Pi = R^2\pi$), а тако се смањује и густина јединичне њене површине (специфична густина), опада јој амплитуда иако таласна дужина остаје иста (које постоје само у виђењу њеног релативног посматрача). Површина дате сфере тренутна је, што значи да је квантно спрегнута, те се сва испоручује у евентуалној интеракцији.

Вероватноћа $p \in (0,1)$ интеракције јединичне површине сфере опада са квадратом полупречника R. Она носи информацију[6] $h = -\ln p \in (0, +\infty)$, количину за коју претпостављамо да не зависи од избора дела њене јединичне површине. Не догађа се непосредна укупна интеракција другог тела са целом сфером, јер је истовременост сфере њена индивидуална појава. Чак ни наше ноге нису истовремене са главом, јер је светлости потребно неко време да стигне са краја на крај, а камоли да нам делови какве веће сфере могу бити истовремени. Зато опажања делова сфере сматрајмо различитим (дисјунктним) догађајима.

Вероватноћа интеракције јединичне површине сфере има униформну расподелу $p = 1/R^2\pi$, њој одговара Хартлијева информација $h(R) = -\ln(R^2\pi)$ која се мења брзином:

$$\frac{dh}{dt} = -\frac{1}{R}\frac{dR}{dt} = -\frac{c}{R}$$

или једноставније писано $R\dot{h} + c = 0$. Информација дела сфере смањује се са повећањем њеног полупречника, што према начелном минимализму информације значи да се сфера шири.

Иначе, тамо где нема промене вероватноће, нема ни промене Хартлијеве информације, из $\dot{p} = 0$ следи $\dot{h} = 0$, а исто важи за Шенонову. Ово друго видимо из:

$$\dot{s} = \frac{ds}{dt} = -\frac{d}{ds}(p_1 \ln p_1 + \dots + p_n \ln p_n) = -[(\dot{p}_1 \ln p_1 + \dot{p}_1) + \dots + (\dot{p}_n \ln p_n + \dot{p}_n)]$$

Отуда, ако је свако $\dot{p}_\nu = 0$ редом за $\nu = 1, \dots, n$ онда је $\dot{s} = 0$. Дакле, из константности вероватноћа следи константност информације. Овде је $p_1, \dots, p_n$ била нека расподела вероватноћа, а Шенонова информација $s = s(n) = -(p_1 \ln p_1 + \dots + p_n \ln p_n)$ је њихова средња вредност.

[6] Информатичка Теорија I, 05. Хартлијева информација, https://archive.org/details/it_20220709

Приметимо да је Шенонова информација (дела сфере потенцијала) све већа када n расте. Наиме, ако на n делова делимо равномерну расподелу, појединa вероватноћа је $p_j = \frac{1}{n}$, па је Шенонова информација:

$$s(n) = -(p_1 \ln p_1 + \cdots + p_n \ln p_n) =$$

$$-\left(\frac{1}{n}\ln\frac{1}{n} + \cdots + \frac{1}{n}\ln\frac{1}{n}\right) = \frac{1}{n}\ln n + \cdots + \frac{1}{n}\ln n = \ln n$$

а $\ln n < \ln(n+1)$, тј. $s(n) < s(n+1)$. Са друге стране гледано, такође је узорак дела сфере дате, рецимо јединичне, површине све информативнији, јер већа сфера граница је већој неизвесности.

Међутим, удео информације у појединoм делу, сабирку Шенонове информације, све је мањи, јер

$$\frac{1}{n+1}\ln(n+1) < \frac{1}{n}\ln n$$

за свако $n > e = 2.71828...$ Доказ је извод функције $f(x) = \frac{1}{x}\ln x$, који је

$$f'(x) = \frac{1 - \ln x}{x^2} < 0$$

за $\ln x > 1$, тј. $x > e$, када је ова функција опадајућа и важи претходна неједнакост. То би ишло са смањивањем дела површине растуће сфере репрезентативног за комуникацију.

На начин сличних сфера објашњавао сам ширење виртуелни фотона око електрона, иза тестирања дубљег смисла у Фајнмановим дијаграмима[7]: „Да би узрок одбијања електрона Кулоновом силом (пропорционалном наелектрисањима честица а обрнуто пропорционалном квадрату растојања између њих) били виртуелни фотони, не треба их третирати као одскакивања пинг-понг лоптице. Рачун ће бити тачан тек ако фотон замишљамо као виртуелну сферу која се шири из центра (сваког од) електрона." Размене теку у пакетима, испоруком кванта дејства настаје дефицит информације, а због начелног минимализма (привлачности мањка) информације, настаје супротна емисија.

У случају мало општије, информације перцепције $S = x_1y_1 + \cdots + x_ny_n$, која је скаларни производ вектора $\vec{x} = (x_1, \ldots, x_n)$ и $\vec{y} = (y_1, \ldots, y_n)$, постоје даље могућности одржања. На пример, из

$$\dot{S} = (\dot{x}_1y_1 + x_1\dot{y}_1) + \cdots + (\dot{x}_ny_n + x_n\dot{y}_n)$$

може се добити константност у времену, $\dot{S} = 0$, ако је $\dot{x}_\nu y_\nu + x_\nu \dot{y}_\nu = 0$ редом за свако $\nu = 1, \ldots, n$. То значи раст прве варијабле када друга опада и обрнуто, осциловање какво се заиста догађа.

[7] Многострукости, Квантна спрегнутост, http://rvukovic.net/knjige/Mnogostrukosti.pdf

47. Осцилације

Периодичне појаве су начин да се одржавају количине, а да системи који немају „памћење" стару вест опет виде као нову. За системе са слабим памћењем излаз је слаба периодичност, а у случају појава јачег памћења ту су „понављања" на увек довољним изузецима. Такав је свет засниван на информацијама чија суштина је неизвесност.

Као што знамо, диференцијал производа је $d(xy) = (dx)y + x(dy)$. Он може бити нула, а да се те две величине стално мењају, $(dx)y = -x(dy)$, када повећање једне прати одговарајуће смањење друге тако да трајно буде $d(xy) = 0$. Када имамо спрегу $S = x_1y_1 + \dots + x_ny_n$, компоненти низова $\vec{x} = (x_1, \dots, x_n)$ и $\vec{y} = (y_1, \dots, y_n)$, исто је могуће са сваким паром фактора у њеним сабирцима. Ово показује значај информације перцепције у теорији информације каквом је развијам.

Периодична непрекидна таласна функција $\psi = \psi(x_1, \dots, x_n)$ задовољава таласну једначину

$$\frac{\partial^2\psi}{\partial t^2} = c^2\left(\frac{\partial^2\psi}{\partial x_1^2} + \dots + \frac{\partial^2\psi}{\partial x_n^2}\right)$$

где је c^2 реална константа. У физици t је време, па надовезујући се на разматрања из претходних наслова, да се садашњост кроз простор-време васионе креће брзином светлости, тада је c брзина светлости. Такво време постаје просторна координата $\tau = ict$, где за имагинарну јединицу важи једнакост $i^2 = -1$ и у физичком 4-Д простор-времену можемо писати

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{\partial^2\psi}{\partial \tau^2} = 0$$

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = 0$$

а то је таласна једначина која следи из Максвелових радова о електромагнетизму, где је $\frac{1}{c^2} = \mu_0\varepsilon_0$, а $\mu_0 = 4\pi \times 10^{-7}$ N/A^2 магнетна, а $\varepsilon_0 = 8{,}854 \times 10^{-12}$ F/m електрична константа пермеабилности вакуума[8].

Симетрија просторних и временских координата већ је примећена у физици, путем Максвелових једначина, па теоријом релативности, али она никада није довољно експлицитно изражена као у открићу 6-Д простор-времена „теорије информације" (Dimensions). Међутим, ни то није довољно, или кажимо да је прихватање нових димензија времена отварање путева у псеудо-реалности, а то даље у тумачење комплексних бројева као врсте величина „стварности" (Solenoid).

На случајан начин честица бира више могућности, од којих ће нека бити да она остаје у „нашој" реалности. Међутим, њен другачији избор „постојаће" и у некој од псеудо-реалности, из које она даље може опет бирати, а ако је неки од тих избора полазна реалност настаје појава коју називам „заобилажење" (Bypass). Алгебарски, то би значило да било која па и реална функција, $f(x) \in \mathbb{R}$ за свако $x \in \mathbb{R}$, која би била неко комплексно решење датог задатка, $f(z) \in \mathbb{C}$ за $z \in \mathbb{C}$, сабрана са себи коњугованом вредношћу $f^*(z)$, давала би реалну вредност $f + f^* \in \mathbb{R}$.

[8] Простор-Време, Пример 1.4.3. http://rvukovic.net/knjige/Prostor-Vreme.pdf

Има и други начини да се добије реалан број алгебарским операцијама са комплексним, наравно, а и многих од њих који би интерпретирали физичке појаве у процесима „заобилажења". Новост је (моје) теорије информације објава да би такав рачун имао неког физичког смисла.

За представљање решења произвољне полиномске једначине ($c_0 + c_1 z + c_2 z^2 + \dots + c_n z^n = 0$) са реалним коефицијентима нису довољни реални бројеви, али са комплексним се могу исказати сви корени таквих, чак и ако су јој коефицијенти комплексни. Та универзалност комплексних бројева у сличним ситуацијама даје прилику да, бар привремено, заокружимо причу о простору, времену и материји на поменутих 3+3 димензије простора и времена, чија структура је информација, а њена дубља суштина неизвесност.

Тако долазимо до универзалног комплексног броја, којег пишемо на различите начине:

$$e^{i\varphi} = \exp(i\varphi) = \operatorname{cis}\varphi = \cos\varphi + i\sin\varphi$$

Његов модуо, интензитет је један $|e^{i\varphi}| = 1$, а множење њиме произвољног броја

$$z = a + ib = re^{i\theta},$$

где је

$$r = |z| = \sqrt{a^2 + b^2} \text{ и } \operatorname{tg}\theta = \frac{b}{a},$$

представља ротацију датог у комплексној равни око њеног исходишта за угао φ. Наиме, у најопштијем случају $z = re^{i\alpha}$ (за $\alpha = 0, \pm 2\pi, \pm 4\pi, \dots$ тај број је реалан):

$$ze^{i\varphi} = re^{i\theta}e^{i\varphi} = re^{i(\theta+\varphi)}.$$

Међутим, ротација је таква изометријска трансформација (која чува удаљености тачака) која може заменити све остале (транслације, рефлексије, ротације).

Тим двема „универзалностима" комплексних бројева, у исказивању решења једначина и замена у описивању изометријских трансформација, треба додати и физичке процесе који изражавају закон одржања. У просторима вероватноћа, удаљености су величине, какве су посебно и обзервабле. То последње на свој начин примећено је у квантној физици.

Шлаг на торту овде пробране ретроспективе о осцилацијама су логаритми комплексних бројева

$$\operatorname{Log} z = \ln r + i\theta$$

који не мењају вредност након пуне ротације, за угао $2k\pi$, где је k произвољан цели број. Дакле, када угао θ заменимо углом $\theta + 2k\pi$ добићемо исту вредност овог логаритма, па је функција $\operatorname{Log} z$ периодична. Проширењем Хартлијеве информације ($h = \log z$), са изворно z природних бројева, па на позитивне реалне и коначно на комплексне бројеве добијамо начин за опис света каквим он јесте, структуриран периодичним информацијама. У Причама о информацији (2.1 Фреквенција) писао сам један популаран, можда необичан прилог о осцилацијама који препоручујем, ако вам је овај текст био претежак.

48. Лапласова трансформација

Лапласова трансформација је једна од разних математици познатих трансформација интеграла, уз Фуријеову, Пуасонову, Мелинову и сличне. Њоме решавамо диференцијалне једначине рецимо, трансформишући их у алгебарске чији резултат затим враћамо инвезијом. Рачунске поступке овм трансформацијом сводимо на алгебарске, пресликавамо величине које су функције времена t у функције комплексне варијабле z.

Нека је $\boldsymbol{F}$ скуп реалних или комплексних функција f реалне променљиве t, $0 < t < +\infty$, које су непрекидне у скоро свим (сем можда у изолованим) тачкама домена и нека постоји $\int_a^b |f(t)|\, dt$ за свако $0 < a < b < \infty$. Функција f не мора бити ограничена за све променљиве, али за бар један реалан или комплексан број z постоји граница $\lim_{T\to\infty} \int_0^T e^{-zt} f(t)\, dt$.

Дефиниција 1. За функцију $f \in \boldsymbol{F}$**,** интеграл $\mathcal{L}\{f(t)\} = \int_0^\infty e^{-zt} f(t)\, dt$ називамо „Лапласовим интегралом“ те функције, а $F(z) = \mathcal{L}\{f(t)\}$ називамо „Лапласовом трансформацијом“. ▨

Кажемо да је интеграл „апсолутно конвергентан“ ако постоји $\lim_{T\to\infty} \int_0^T |e^{-zt} f(t)|\, dt$ тако да за све $T' > T$ важи $\left|\int_T^{T\prime} e^{-zt} f(t)\, dt\right| \le \int_T^{T\prime} |e^{-zt} f(t)|\, dt \to 0$ када $t \to 0$. Функција је „експоненцијално оганичена“ ако постоје константе $a > 0$ и $b \in \mathbb{R}$ такве да за неко $t_0 \ge 0$ вреди $|f(t)| \le a e^{bt}$ када $t \ge t_0$. Функција f је део по део непрекидна на интервалу $[0, \infty)$ ако постоји $\lim_{t\to 0^+} f(t) = f(0^+)$ и непрекидна је на сваком коначном интервалу (r_1, r_2) осим у евентуално коначно много тачака у којима има прекиде.

Теорема множења (Борелова теорема) говори о „конволуцији“ функције $f(x)$ и $g(x)$ означавану са $f(t) * g(t)$ за коју важи $\mathcal{L}\{f * g\} = \mathcal{L}\{f\}\mathcal{L}\{g\}$.

48.1. Основне особине

Са операцијама сабирања и множења комплексним бројем, скуп $\boldsymbol{F}$ има структуру векторског простора над телом комплексних бројева. То следи из линеарности:

$$\mathcal{L}\{c_1 f_1 + c_2 f_2\} = \int_0^\infty e^{-zt}[c_1 f_1(t) + c_2 f_2(t)]\, dt = c_1 \int_0^\infty e^{-zt} f_1(t)\, dt + c_2 \int_0^\infty e^{-zt} f_2(t)\, dt$$

Важи „теорема сличности“ која каже да је

$$\mathcal{L}\{f(at)\} = \frac{1}{a} f\left(\frac{z}{a}\right)$$

и „теорема померања“, једнакост

$$F(z - a) = \mathcal{L}\{e^{at} f(t)\}$$

Обе важе у условима када дати интергал постоји (рецимо експоненцијалне ограничености). Њих је лако доказати.

Лапласова трансформација се у математици користи за анализирање линеарних, временски непромењивих система, какве имамо код електричних кола, хармонијских осцилатора, оптичких

уређаја и механичких система. Има примене у решавању диференцијалних једначина и теорији вероватноће. Због овако велике примењивости неће бити чудно да ускоро откријемо и велику примењивост Лапласове трансформације у (будућој) теорији информације.

Пример 1. Показати да је: $\mathcal{L}[c] = \frac{c}{z}$, $\mathrm{Re}(z) > 0, c \in \mathbb{C}$.

Решење: $\mathcal{L}[c] = F(z) = \int_0^\infty e^{-zt} c\, dt = \frac{c}{z}\int_0^\infty (-e^{-zt})\, d(-zt) = \frac{c}{z}(-e^{-zt})_0^\infty = \frac{c}{z}$. ❑

Пример 2. Ако је $f(t) = e^{bt}$, $b \in \mathbb{R}$, тада је $\mathcal{L}[e^{bt}] = \frac{1}{z-b}$.

Решење: $\mathcal{L}\{e^{bt}\} = \int_{t=0}^\infty e^{-zt} e^{bt}\, dt = \int_{t=0}^\infty e^{(b-z)t}\, dt = \frac{1}{b-z}\int_{x=0}^\infty e^x\, dx = \frac{1}{b-z} e^x|_{x=0}^\infty = \frac{1}{z-b}$. ❑

Ево још неколико примера Лапласове трансформације, за $n = 1, 2, 3, \dots$, $b > 0$, Гама функцију, косинус, синус и хиперболни косинус и синус. Доказе можете наћи на много места.

$f(t)$	$\mathcal{L}\{f\}$		$f(t)$	$\mathcal{L}\{f\}$
1	$\frac{1}{z}$		e^{bt}	$\frac{1}{z-b}$
t	$\frac{1}{z^2}$		$\cos \alpha t$	$\frac{z}{z^2+\alpha^2}$
t^2	$\frac{2!}{z^3}$		$\sin \alpha t$	$\frac{\alpha}{z^2+\alpha^2}$
t^n	$\frac{n!}{z^{n+1}}$		$\mathrm{ch}\, \alpha t$	$\frac{z}{z^2-\alpha^2}$
t^b	$\frac{\Gamma(b+1)}{z^{b+1}}$		$\mathrm{sh}\, \alpha t$	$\frac{\alpha}{z^2-\alpha^2}$

Следећи пример је заправо позната „теорема о помаку“. Зато наводим и доказ, иако није тежак.

Пример 3. Како је $F(z) = \int_0^\infty e^{-zt} f(t) dt$, за биће:

$$F(z-b) = \int_0^\infty e^{-(z-b)t} f(t)\, dt = \int_0^\infty e^{-zt}[e^{bt} f(t)]\, dt = \mathcal{L}\{e^{bt} f(t)\}$$

за $z - b > 0$. ❑

Последњим пример говори и о „чисто временском кашњењу“. Две функције истог облика, $f(t)$ и $f(t - t_0)$, где друга касни за првом за време t_0, могуће је дефинисати са $\mathcal{L}[f(t-t_0)] = e^{-zt_0}F(z)$. Чисто временско кашњење има рецимо трофазни систем наизменичних струја.

Користећи исти 3. пример, лако допуњавамо претходну табелу, налазимо:

- ако је $f(t) = e^{bt} t^n$, онда је $\mathcal{L}\{f\} = \frac{n!}{(z-b)^{n+1}}$;
- ако је $f(t) = e^{bt} \cos \alpha t$, онда је $\mathcal{L}\{f\} = \frac{z-b}{(z-b)^2+\alpha^2}$;

- ako je $f(t) = e^{bt}\sin\alpha t$, онда је $\mathcal{L}\{f\} = \frac{\alpha}{(z-b)^2+\alpha^2}$.

Када је функција $f(t)$ део по део непрекидна на интервалу $[0,\infty)$ и експоненцијално ограничена, тада њена Лапласова трансформација постоји. То каже следећи пример, теорема егзистенције.

Пример 4. Када је функција $f(t)$ по деловима непрекидна за сваки коначни интервал где је $t > 0$ и ако је $|f(t)| \le Me^{bt}$ за све такве променљиве и неке константе b и M, тада постоји Лапласова трансформација дате функције за све $z > b$. Наиме:

$$|\mathcal{L}\{f\}| = \left|\int_0^\infty e^{-zt} f(t)\,dt\right| \le \int_0^\infty e^{-zt}|f(t)|\,dt \le \int_0^\infty e^{-zt}\,Me^{bt}dt = M\int_0^\infty e^{-(z-b)t}\,dt = \frac{M}{z-b}$$

што је конвергенција. ❑

Две функције које имају једнаку Лапласову трансформацију потпуно су идентичне. Друга згода са Лапласовим трансформацијама је што множење и дељење $F(z) = \mathcal{L}\{f\}$ одговара диференцирању и интегрирању функције $f(t)$. На тај начин се сложене операције инфинитезималног рачуна своде на једноставне алгебарске.

48.2. Инфинитезимале

У условима егзистенције (f је непрекидна за све променљиве $t \ge 0$, и $|f(t)| \le Me^{bt}$ за неке константе b и M), важиће једнакости за изводе функције $f(t)$ редом:

$$\mathcal{L}\{f'\} = z\mathcal{L}\{f\} - f(0)$$

$$\mathcal{L}\{f''\} = z^2\mathcal{L}\{f\} - zf(0) - f'(0)$$

$$\mathcal{L}\{f'''\} = z^3\mathcal{L}\{f\} - z^2 f(0) - zf'(0) - f''(0)$$

...

$$\mathcal{L}\{f^{(n)}\} = z^n\mathcal{L}\{f\} - z^{n-1}f(0) - z^{n-2}f'(0) - \cdots - f^{(n-1)}(0)$$

Наиме,

$$\mathcal{L}\{f'\} = \int_0^\infty e^{-zt}\,f'(t)\,dt = [e^{-zt}f(t)]_0^\infty + z\int_0^\infty e^{-zt}f(t)\,dt$$

Тиме је доказана наведена Лапласова трансформација за први извод. Уврштавањем те вредности у другу деривацију добијамо:

$$\mathcal{L}\{f''\} = z\mathcal{L}\{f'\} - f'(0) =$$

$$= z[\mathcal{L}\{f\} - f(0)] - f'(0)$$

$$= z^2\mathcal{L}\{f\} - zf(0) - f'(0)$$

Тиме је доказана ЛТ за други извод. Слично добијамо трећу и остале деривације.

Такође обрнуто, у условима егзистенције, важиће једнакост

$$\mathcal{L}\left\{\int_0^\infty f(s)\,ds\right\} = \frac{1}{z}\mathcal{L}\{f(t)\}$$

Наиме, за $z > 0$ и $b > 0$ претпоставимо да је интеграл $g(t) = \int_0^t f(s)\,ds$ непрекидна функција. Из:

$$|g(t)| \le \int_0^t |f(s)|\,ds \le M\int_0^t e^{bs}\,ds = \frac{M}{b}\left(e^{bs} - 1\right)$$

Следи ограниченост и $g'(t) = f(t)$, осим у тачкама где $f(t)$ има прекид. Отуда, према претходним једнакостима за изводе:

$$\mathcal{L}\{f(t)\} = \mathcal{L}\{g'(t)\} = z\mathcal{L}\{g(t)\} - g(0)$$

па како је $g(0) = 0$ добијамо

$$\mathcal{L}\{g\} = \frac{1}{z}\mathcal{L}\{f\}$$

а отуда горња једнакост за ЛТ интеграла. Примену ових формула видимо у следећем примеру.

Пример 5. Решавамо диференцијалну једначину другог реда са константним коефицијентима

$$\ddot{y}(t) + a\dot{y}(t) + by(t) = ct$$

уз почетне услове $y(0) = -1$ и $\dot{y}(0) = 2$. Узимамо Лапласову трансформацију леве и десне стране једнакости и на основу линеарности добијамо:

$$\mathcal{L}\{y''\} + a\mathcal{L}\{y'\} + b\mathcal{L}\{y\} = \mathcal{L}\{ct\}$$

$$z^2Y - zf(0) - f'(0) + a[zY - f(0)] + bY = \frac{c}{z^2}$$

$$(z^2 + az + b)Y + z + a - 2 = \frac{c}{z^2}$$

где је $Y = \mathcal{L}\{f\}$, $f(0) = -1$ и $f'(0) = 2$, а из таблице узимамо $\mathcal{L}\{ct\} = c\mathcal{L}\{t\} = c/z^2$. Отуда

$$Y(z) = \frac{-z^3 - (a-2)z^2 + c}{z^2(z^2 + az + b)}$$

Добили смо алгебарску једначину која се назива „помоћна једначина“. Она се затим разлаже у „парцијалне разломке“ и из таблице налазе решења. ❑

Да би разумели о чему се ради у завршници 5. примера, ставимо $a = -10$, $b = 9$ и $c = 5$, ради поједностављивања и довршимо рачун:

$$Y(z) = \frac{-z^3 + 12z^2 + 5}{z^2(z^2 - 10z + 9)} = \frac{5}{z^2(z-9)(z-1)} + \frac{12 - z}{(z-9)(z-1)} =$$

$$= \frac{5}{9z^2} + \frac{50}{81z} + \frac{31}{81(z-9)} - \frac{2}{z-1}$$

$$= \frac{50}{81} + \frac{5}{9}t + \frac{31}{81}e^{9t} - 2e^{t}$$

У литератури има довољно оваквих решења, намењених увежбавању. Приметимо да диференцијалне једначине говоре о улазним и излазним величинама.

48.3. Инверзна трансформација

Нека је $\mathcal{L}\{f(t)\} = F(z)$. Решење једначине $F(z) = \int_0^t e^{-zt} f(t)\,dt$ где је $F(t)$ познато, назива се инверзна Лапласова трансформација и пише $f(t) = \mathcal{L}^{-1}\{F(z)\}$. Основне особине инверзне Лапласове трансформације су:

1. $\mathcal{L}^{-1}\{aF(z) + bG(z)\} = af(t) + bg(t)$;
2. $\mathcal{L}^{-1}\{F(cz)\} = \frac{1}{c} f\left(\frac{t}{c}\right)$;
3. $\mathcal{L}^{-1}\{F(z-b)\} = e^{bt} f(t)$;
4. $\mathcal{L}^{-1}\left\{F^{(n)}(z)\right\} = (-1)^n t^n f(t)$;
5. $\mathcal{L}^{-1}\{F(z)G(z)\} = f(t) * g(t)$.

Пример 9. Одређујемо инверзну Лапласову трансформацију функције

$$F(z) = \frac{3z - 14}{4z^2 + 12z + 10}$$

Дату функцију трансформишемо у облик

$$F(z) = \frac{3\left(z + \frac{3}{2}\right) - \frac{37}{6}}{4\left(z + \frac{3}{2}\right)^2 + \frac{1}{4}}$$

погодан за читање решења из таблице. Даље је:

$$f(t) = \mathcal{L}^{-1}\{F(z)\} = \frac{3}{4}\mathcal{L}^{-1}\left\{\frac{3\left(z + \frac{3}{2}\right) - \frac{37}{6}}{4\left(z + \frac{3}{2}\right)^2 + \frac{1}{4}}\right\} = \frac{3}{4}e^{-\frac{3}{2}t}\mathcal{L}^{-1}\left\{\frac{z - \frac{37}{6}}{z^2 + \frac{1}{4}}\right\} =$$

$$= \frac{3}{4}e^{-\frac{3}{2}t}\mathcal{L}^{-1}\left\{\frac{z}{z^2 + \frac{1}{4}}\right\} - \frac{3}{4}\frac{37}{6}e^{-\frac{3}{2}t}2\mathcal{L}^{-1}\left\{\frac{\frac{1}{2}}{z^2 + \frac{1}{4}}\right\}$$

$$= \frac{3}{4}e^{-\frac{3}{2}t}\left(\cos\frac{t}{2} - \frac{37}{3}\sin\frac{t}{2}\right)$$

То је функција $f(t)$ чија је Лапласова трансфомација дата функција $F(z)$. ❑

Лапласова трансформација је „једнозначна“. То значи да задата функција комплексне варијабле, рецимо $F(z)$, која је Лапласова трансформација неке временске функције $f(t)$ и уједно неке друге временске функције $g(t)$, онда је разлика између функција $f(t)$ и $g(t)$ занемарљива. Искључујући ове тривијалне случајеве, ова једнозначност гарантује да је нека временска функција јединствено одређена својом Лаплацеовом трансформацијом.

Из $F(z) = \mathcal{L}\{f(t)\}$ следи $f(t) = \mathcal{L}^{-1}\{F(z)\}$, другим речима функција $f(t)$ је инверзна Лапласова трансформација од $F(z)$.

48.4. Пригушивање

Пример 6. Једначина $m\dot{v}(t) + bv(t) = f(t)$ говори, лево од једнакости о траженим брзинама, на основу улазне силе $f(t)$ на десној страни. Ставимо ли да за масу $m = 1$, $b = 2$ и $y = v(t)$, то ће бити „нема" диференцијална једначина првог реда са константним коефицијентима

$$y'(t) + 2y(t) = f(t)$$

где улаз може бити и струја, а излаз напон и слично. Као и у 5. примеру, заменама ЛТ, налазимо овај пут помоћну једначину

$$zY(z) + 2 + 2Y(z) = \frac{1}{z} - \frac{e^{-z}}{z}$$

коју затим треба разлагати у парцијалне разломке и из таблице (или израчунавањем обрнуте ЛТ) наћи решења. ❑

Оно што Лапласова трансформација ради са улазном величином је пригушивање током времена, тако мењајући излазне резултате. Како она пригушује улаз, упоредо Фуријеовој трансформацији (која разлаже функцију времена у фреквенције које га чине), визуелно је приказано у занимљивом прилогу What does the Laplace Transform tell us. Она је таква нарочито згодна за опис смањивавња густине вероватноће, или амплитуда, временом или повећањем полупречника виртуелних сфера које смо разматрали у претходним насловима.

Пример 7. (Пригушена синусна функција) Нека за реални део комплексног броја важи $\mathrm{Re}(z) > -a$, за неки реалан позитиван број a и нека је:

$$f(t) = e^{-at}\sin\omega t\, x(t) = e^{-at}\left(\frac{e^{i\omega t} - e^{-i\omega t}}{2i}\right)x(t)$$

Тада је Лапласова трансформација:

$$F(z) = \mathcal{L}\{f\} = \mathcal{L}\{e^{-at}\sin\omega t\, x(t)\} = \mathcal{L}\left\{e^{-at}\left(\frac{e^{i\omega t} - e^{-i\omega t}}{2i}\right)x(t)\right\} =$$

$$= \frac{1}{2i}\mathcal{L}\{(e^{-at}e^{i\omega t} - e^{-at}e^{-i\omega t})x(t)\} = \frac{1}{2i}\mathcal{L}\{[e^{-(a-i\omega)t} - e^{-(a+i\omega)t}]x(t)\}$$

$$= \frac{1}{2i}\mathcal{L}\{e^{-(a-i\omega)t}x(t)\} - \frac{1}{2i}\mathcal{L}\{e^{-(a+i\omega)t}x(t)\}$$

$$= \frac{1}{2i}\left[\frac{1}{(z+a) - i\omega} - \frac{1}{(z+a) + i\omega}\right] = \frac{\omega}{(z+a)^2 + \omega^2}$$

Област конвергенције ЛТ је $\mathrm{Re}(z) > -a$, дате пригушене синусне функције. ❑

Пример 8. (Пригушена косинусна функција) За реални део комплексног броја важи $\mathrm{Re}(z) > -a$, за неки реалан позитиван број a и дато је:

$$f(t) = e^{-at}\cos\omega t\, x(t) = e^{-at}\left(\frac{e^{i\omega t} + e^{-i\omega t}}{2i}\right)x(t)$$

Помоћу Лапласове трансформације добијамо:

$$F(z) = \mathcal{L}\{f\} = \mathcal{L}\{e^{-at}\sin\omega t\, x(t)\} = \mathcal{L}\left\{e^{-at}\left(\frac{e^{i\omega t} + e^{-i\omega t}}{2}\right)x(t)\right\} =$$

$$= \frac{1}{2}\mathcal{L}\{(e^{-at}e^{i\omega t} + e^{-at}e^{-i\omega t})x(t)\} = \frac{1}{2}\mathcal{L}\{[e^{-(a-i\omega)t} + e^{-(a+i\omega)t}]x(t)\}$$

$$= \frac{1}{2}\left[\frac{1}{z + (a - i\omega)} + \frac{1}{z + (a + i\omega)}\right] = \frac{z + a}{(z + a)^2 + \omega^2}$$

Област конвергенције ЛТ такође је $\mathrm{Re}(z) > -a$, сада пригушене косинусне функције. ❑

Обзиром да је $e^{i\omega t} = \cos\omega t + i\sin\omega t$, то за функцију

$$f(t) = e^{-at}e^{i\omega t}x(t) = e^{-at}(\cos\omega t + i\sin\omega t)x(t)$$

Лапласову трансформацију добијамо као збир примера 7-8, односно:

$$\mathcal{L}\{f\} = \mathcal{L}\{e^{-at}e^{i\omega t}x(t)\} = \frac{z + (a + i\omega)}{(z + a)^2 + \omega^2}$$

Област конвергенције Лапласове трансформације ове пригушене функције је $\mathrm{Re}(z) > -a$. Када угао неограничено расте, $\omega \to \infty$, тада:

$$\mathcal{L}\{f\} = \frac{\frac{z}{\omega} + \left(\frac{a}{\omega} + i\right)}{\frac{(z + a)^2}{\omega} + \omega} \to \frac{i}{\omega} \to 0$$

Она представља и начин пригушивања таласне функције слободне честице квантне механике, из чега лако видимо бар две последице. Смањивање вредности амплитуде приближно одговара опису опадања потенцијала [44-45.] виртуелних сфера и, при исчезавању, његовим преласком у чисто имагинарни домен.

49. Фуријеова трансформација

Фуријеова трансформација је математичка операција разлагања периодичне функције на њене „спектралне компоненте“ ради једноставније анализе. Она узима шаблон заснован на времену, мери могуће циклусе и враћа укупан „рецепт циклуса“, тј. њихове амплитуде, помаке и брзине ротације. Уопште, математичка трансформација је промена перспективе. У том смислу Фуријеова трансформација мења нашу перспективу од „потрошача“ до „произвођача“, претварајући „шта имам“ у „како је направљен“, за дату смешу она тражи рецепт.

Са становишта инжењерства, Фуријеова трансформација је „алатка“ за одређивање доминантних фреквенција у сигналима вибрација. Када доминантна фреквенција сигнала одговара природној фреквенцији структуре, вибрације које се јављају могу се појачати услед резонанце. Ово се може догодити до те мере да сруши структуру.

Дефиниција 1. Дискретна Фуријеова трансформација претвара низ $x_1, \dots, x_N$ датих N комплексних бројева у низ N комплексних бројева

$$X_k = \sum_{n=0}^{N-1} x_n e^{2\pi ikn/N}$$

за $0 \le k \le N-1$. ▨

Иначе је за имагинарну јединицу $i^2 = -1$, а остале ознаке су:

- N — број узорака,
- n — редни број текућег узорка,
- x_n — величина сигнала у n-том тренутку,
- k — текућа фреквенција (од 0 Hz до $N-1$ Hz),
- X_k — резултат дискретне Фуријеове трансформације (амплитуда и фаза).

Излаз X_k је комплексан број који кодира амплитуду и фазу синусоидног таласа са фреквенцијом k/N циклуса у јединици времена, на основу Ојлерове формуле

$$\exp(2\pi ikn/N) = \cos(2\pi kn/N) + i\sin(2\pi kn/N)$$

Израчунавањем X_k проналазе се коефицијенти апроксимације сигнала линеарном комбинацијом таквих таласа. Пошто сваки талас има цео број N циклуса током јединице времена, апроксимација ће бити периодична са периодом N.

Дефиниција 2. Инверзна дискретна Фуријеова трансформација враћа низ $X_1, \dots, X_N$ нађених N комплексних бројева у низ N оригиналних комплексних бројева

$$x_n = \frac{1}{N}\sum_{k=0}^{N-1} X_k e^{2\pi ikn/N}$$

за $0 \le n \le N-1$. ▨

Сви таласни облици, без обзира шта пишемо или посматрамо у васиони, заправо су само збир једноставних синусоида различитих фреквенција. Отуда могућност да се функције тако масовно и

успешно подвргавају Фуријеовој трансформацији. Штавише, представљање скоро сваког облика линије графа (трајекторије честице) помоћу произвољних таласних облика сугерише нам да су све те линије заправо њени облици, односно да линије и немају истакнуте микро облике.

Приметимо да својом формом Фуријеова трансформација подсећа на „информацију перцепције“. Она је скаларни производ вектора (низова) опажаја субјекта и објекта који дефинишемо са

$$\vec{a}\cdot\vec{b}=\sum_{k=1}^{n}a_k b_k$$

Иста је форма дискретне Фуријеове трансформације из чега произилазе и неке (апстрактне) исте изведбе. Међутим, ова се дискретна форма лако преноси у одговарајуће непрекидне облике.

Дефиниција 3. Непрекидна Фуријеова трансформација функције $f(t)$ је

$$F(\omega)=\mathcal{F}\{f\}=\int_{-\infty}^{+\infty}e^{-i\omega t}f(t)\,dt$$

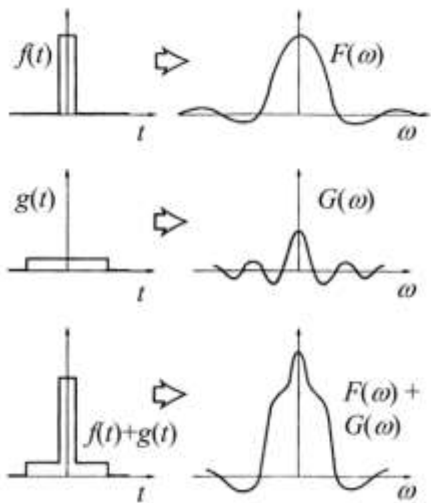
где је $\omega=2\pi\varphi$ обично неки угао и t је време. Инверзна Фуријеова трансформација овој је

$$f(t)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{i\omega t}F(\omega)\,d\omega.$$

Функције $f(t)$ и $F(\omega)$, односно $\mathcal{F}\{f\}$, интеграбилне су комплексне величине. Модуо ове друге је тзв. спектрална густина амплитуда, а аргумент спектрална густина фаза. ▨

Постоји неколико нама интересантних теорема Фуријеових трансформација које ћу овде укратко прелистати и доказати за континуалне случајеве.

Теорема 1. (Линеарности) $\mathcal{F}\{af+bg\}=a\mathcal{F}\{f\}+b\mathcal{F}\{g\}$, где су a и b константе.



Доказ:

$$\mathcal{F}\{af+bg\}(\omega)=\int_{*\infty}^{+\infty}[af(t)+bg(t)]e^{-2\pi i\omega t}\,dt=$$

$$=a\int_{*\infty}^{+\infty}f(t)e^{-2\pi i\omega t}\,dt+b\int_{*\infty}^{+\infty}g(t)e^{-2\pi i\omega t}\,dt$$

$$=aF(z)+bG(z)=a\mathcal{F}\{f\}+b\mathcal{F}\{g\}.$$

Дакле, $\mathcal{F}\{af+bg\}(\omega)=a\mathcal{F}\{f\}+b\mathcal{F}\{g\}$, а то је оно што је требало доказати. ■

На слици лево видимо да ова, теорема линеарности, има примену у проучавању интерференције таласа.

Фуријеова трансформација садржи исте информације као и оригинална функција, она је додатни начин представљања сигнала (у фреквенцијском, а не у временском домену).

Теорема 2. (Померања) $\mathcal{F}\{f(t-x)\}(\omega) = e^{-2\pi i\omega x}\mathcal{F}\{f\}$.

Доказ:

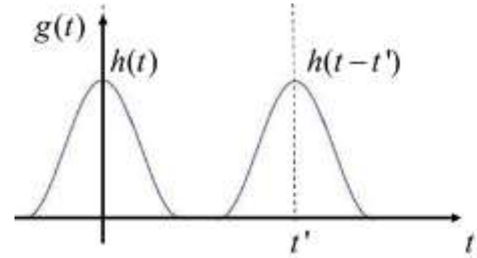
$$\mathcal{F}\{f(t-x)\}(\omega) = \int_{-\infty}^{+\infty} f(t-x)e^{-2\pi i\omega t}\left(e^{2\pi i\omega x}e^{-2\pi i\omega x}\right)dt =$$

$$= e^{-2\pi i\omega x}\int_{-\infty}^{+\infty} f(t-x)e^{-2\pi i\omega(t-x)}\,dt = e^{-2\pi i\omega x}\int_{-\infty}^{+\infty} f(y)e^{-2\pi i\omega y}\,dy$$

$$= e^{-2\pi i\omega x}F(z) = e^{-2\pi i\omega x}\mathcal{F}\{f\}$$

Дакле, $\mathcal{F}\{f(t-x)\}(\omega) = e^{-2\pi i\omega x}\mathcal{F}\{f\}$, што је и требало доказати. ■

Теорема померања говори нам о временски независној форми импулсног одзива. Слика десно показује систем инваријантан на линеарни помак. Одговор на импулс $h(t) \to h(t-t')$ померен из тренутка t за t' даје исти облик импулсног одзива.



49.1 Примери ДФТ

Следе класични примери дискретне Фуријеове трансформације.

Пример 1. Нека је $x_0 = 1$, $x_1 = x_2 = \cdots = x_{N-1}$. Тада је:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi ikn}{N}} = x_0 + x_1 e^{-\frac{2\pi ik}{N}} + x_2 e^{-\frac{2\pi ik2}{N}} + \cdots + x_{N-1}e^{-\frac{2\pi ik(N-1)}{N}} = 1,$$

за свако $k = 0,1,2,\dots,N-1$. ❑

Из овог примера, за инверзно

$$x_n = \frac{1}{N}\sum_{k=0}^{N-1} X_k e^{\frac{2\pi ikn}{N}}$$

следи $x_0 = \frac{1}{N}\cdot N = 1$, док за остале индексе, $n \neq 0$, имамо:

$$x_n = \frac{1}{N}\sum_{k=0}^{N-1}\left(e^{\frac{2\pi in}{N}}\right)^k = \frac{1}{N}\frac{\left(e^{\frac{2\pi in}{N}}\right)^N - 1}{e^{\frac{2\pi in}{N}} - 1} = 0.$$

Дискретна Фуријеова трансформација у 1. примеру „шиљак“, из почетног тренутка, рашчлањује на збир таласа (овде једнако пондерисаних) који интерферирају, међусобно се поништавајући у свим тренуцима осим почетног. То се још боље види када шиљак преместимо у следећи тренутак, као у другом примеру.

Пример 2. Нека је $x_1 = 1$, а $x_0 = x_2 = \cdots = x_{N-1} = 0$. Тада је:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi ikn}{N}} = x_0 + x_1 e^{-\frac{2\pi ik}{N}} + x_2 e^{-\frac{2\pi ik2}{N}} + \cdots + x_{N-1} e^{-\frac{2\pi ik(N-1)}{N}} = e^{-\frac{2\pi ik}{N}},$$

редом за $k = 0,1,2,\dots,N-1$. Ако је, рецимо $N = 4$, добијамо редом:

$$X_0 = e^{-2\pi i\cdot\frac{0}{4}} = e^0 = 1$$

$$X_1 = e^{-2\pi i\cdot 1/4} = \cos\left(-\frac{2\pi}{4}\right) + i\sin\left(-\frac{2\pi}{4}\right) = 0 - i$$

$$X_2 = e^{-2\pi i\cdot 2/4} = \cos(-\pi) + i\sin(-\pi) = -1 + i\cdot 0$$

$$X_3 = e^{-2\pi i\cdot 3/4} = \cos\left(-\frac{3\pi}{2}\right) + i\sin\left(-\frac{3\pi}{2}\right) = 0 + i$$

па је $(X_k) = (1, -i, -1, i)$. ❑

Из 2. примера израчунавамо инверзно, оригинале на основу копија:

$$x_n = \frac{1}{N}\sum_{k=0}^{N-1} X_k e^{\frac{2\pi ikn}{N}} = \frac{1}{4}\left(X_0 + X_1 e^{2\pi in\cdot\frac{1}{4}} + X_2 e^{2\pi in\cdot\frac{2}{4}} + X_3 e^{2\pi in\cdot\frac{3}{4}}\right)$$

где фактори $-i, -1, i$ редом представљају фазне помаке таласа за $\frac{3\pi}{2}, \pi, \frac{\pi}{2}$. Даље је:

$$x_n = \frac{1}{4}\left[1 - i\left(\cos\frac{\pi n}{2} + i\sin\frac{\pi n}{2}\right) - (\cos\pi n + i\sin\pi n) + i\left(\cos\frac{3\pi n}{2} + i\sin\frac{3\pi n}{2}\right)\right]$$

Посебно за $n = 1$ биће:

$$x_1 = \frac{1}{4}[1 - i(0 + i\cdot 1) - (-1 + i\cdot 0) + i(0 - i)] = \frac{1}{4}\cdot 4 = 1$$

Међутим, за све остале ($n \neq 1$) је:

$$x_n = \frac{1}{4}\left[i\left(\cos\frac{3\pi n}{2} - \cos\frac{\pi n}{2}\right) - \left(\sin\frac{3\pi n}{2} - \sin\frac{\pi n}{2}\right)\right]$$

$$= \frac{1}{4}\left(-2i\sin\frac{4\pi n}{2}\sin\frac{2\pi n}{2} - 2\sin\frac{2\pi n}{2}\cos\frac{4\pi n}{2}\right) = 0.$$

Уопште ће инверзна трансформација дати почетне вредности, што се види из следећег:

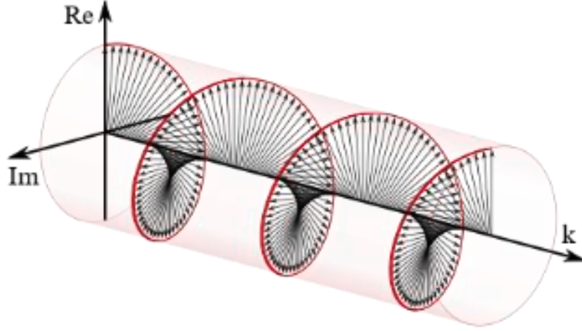
$$x_n = \frac{1}{N}\sum_{k=0}^{N-1} X_k e^{\frac{2\pi ikn}{N}} = x_n = \frac{1}{N}\sum_{k=0}^{N-1}\left(\sum_{m=0}^{N-1} x_n e^{-\frac{2\pi ikm}{N}}\right) e^{\frac{2\pi ikn}{N}} =$$

$$= \frac{1}{N}\sum_{k=0}^{N-1}\sum_{m=0}^{N-1}\left(x_n e^{-\frac{2\pi ikm}{N}}\cdot e^{\frac{2\pi ikn}{N}}\right) = \frac{1}{N}\sum_{k,m=0}^{N-1} x_n e^{\frac{2\pi ik(n-m)}{N}}$$

$$= \frac{1}{N}\sum_{m=0}^{N-1} x_n \sum_{k=0}^{N-1} e^{\frac{2\pi i k(n-m)}{N}} = x_n$$

јер када је $n \neq m$ унутрашњи збир је 0, а са $n = m$ тај збир је N (в. коментар након 1. примера).

Када у 2. примеру за N стављамо све већи број, а и даље једино $x_1 = 1$ је различито од нуле, онда добијамо све више вредности X_k (тачније N таквих), који се могу представљати амплитудама дуж реалне и имагинарне (Re, Im) осе комплексних бројева, док се талас премешта дуж k-осе, као на слици лево. Што је већи број N, учесталост (фреквенција) појављивања ових ротација око k-осе већа је, али са увек истим амплитудама (отклонима) у правцима реалне и имагинарне осе.



Са становишта (моје) „теорије информације“, то има и додатних значења. На пример, када узмемо у обзир многостукост простор-времена (Dimensions), онда су ово ротације наизменично кроз реалност текућег посматрача и његовуу псеудо-реалност, аналогно познатој појави кружне поларизације.

У електродинамици, кружна поларизација (Circular polarization) електромагнетног таласа је стање поларизације у којем, у свакој тачки, електромагнетно поље таласа има константну величину док се ротира константном брзином у равни која је управна на правац таласа.

49.2. Конволуција

Конволуција је израз који нам говори о начину обликовања једне функције (низа) другом, односно она је интеграл (збир) који изражава преклапање једне функције другом. Ту „мешавину“ функција f и g у случају непрекидности прецизно пишемо:

$$\mathcal{F}\{f * g\}(\omega) = \int_{-\infty}^{+\infty} f(t)g(\omega - t)\,dt$$

Сменом промењљиве $y = \omega - t$ следи еквивалентан запис, а отуда комутативност $f * g = g * f$. Из двоструких интеграла приказаних итерацијама налазимо $f * (g * h) = (f * g) * h$, асоцијативност, а из линеарности ингеграла и дистрибутивност $f * (g + h) = (f * g) + (f * h)$.

У дискретном случају конволуције, ради лакшег разумевања, прво посматрајмо два полинома:

$$f(x) = a_0 + a_1 x, \quad g(x) = b_0 + b_1 x + b_2 x^2$$

претпостављајући да су сви коефицијенти који овде нису написани једнаки нули, $a_2 = a_3 = \cdots = 0$ и $b_3 = b_4 = \cdots = 0$. Производ ових полинома је:

$$f(x)g(x) = a_0 b_0 + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0),$$

$$f(x)g(x) = (\vec{a} * \vec{b})_0 + (\vec{a} * \vec{b})_1 x + (\vec{a} * \vec{b})_2 x^2$$

где су

$$(\vec{a} * \vec{b})_k = \sum_{j=0}^{k} a_j b_{k-j}$$

дискретне конволуције. Следећа, теорема конволуције за дискретну Фуријеову трансформацију тврди да у временском домену конволуција одговара множењу по тачкама у фреквенцијском домену, да

$$(g * h) \leftrightarrow G \cdot H$$

где су десно Фуријеове трансформације функција лево. Фуријеова трансформација конволуције је производ две Фуријеове трансформације! Доказ је једноставан и као што се види у наставку доћи ће кроз замену редоследа сабирања из конволуције и дискретне временске трансформације.

Теорема 3. (Конволуције) $\mathcal{F}\{g * h\} = \mathcal{F}\{g\} \cdot \mathcal{F}\{h\}$.

Доказ:

$$\mathcal{F}\{g * h\}(\omega) = \sum_{k=-\infty}^{+\infty} (g * h)_k \, e^{i\omega k} = \sum_{k=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} g_m h_{k-m} e^{-i\omega k} =$$

$$= \sum_{k=-\infty}^{+\infty} g_m \left(\sum_{m=-\infty}^{+\infty} h_{k-m} e^{-i\omega k} \right) = \sum_{k=-\infty}^{+\infty} g_m \left(\sum_{m=-\infty}^{+\infty} H_m \right)$$

$$= \left(\sum_{k=-\infty}^{+\infty} g_k e^{-i\omega k} \right) H(\omega) = G(\omega) H(\omega).$$

На крају је кориштена 2. теорема (померања). Дакле, $\mathcal{F}\{g * h\}(z) = \mathcal{F}\{g\} \cdot \mathcal{F}\{h\}$. ■

Конволуцију $g(t)$ и $h(t)$ кратко означавамо са $g * h$. У применама ове функције често имају сасвим различита значења. Једна као типични сигнал или ток података који траје у недоглед у времену, а друга је функција одговора, обично вршна, поентирана одакле на обе стране пада на нулу.

49.3. Непрекидност

Теорема сличности је у основи ограничена на случај непрекидног времена. Каже да „развлачење“ сигнала за фактор α у временском домену чини „стискање“ његове Фуријеове трансформације за исти фактор у фреквенцијском домену. Ово је веома фундаменталан Фуријеов однос.

Теорема 4. (Сличности) $\mathcal{F}\{f(\alpha t)\}(\omega) = \frac{1}{|\alpha|} F\left(\frac{\omega}{\alpha}\right)$.

Доказ: $\mathcal{F}\{f(\alpha t)\} = \int_{-\infty}^{+\infty} f(\alpha t) e^{-\frac{2\pi i \omega}{\alpha}\alpha t} \, \alpha dt =$

$$= \begin{cases} \dfrac{1}{|\alpha|} \displaystyle\int_{t=-\infty}^{+\infty} f(\alpha t) e^{-2\pi i \frac{\omega}{\alpha}(\alpha t)} \ \alpha dt, & \alpha > 0 \\ \dfrac{1}{|\alpha|} \displaystyle\int_{t=-\infty}^{+\infty} f(-|\alpha| t) e^{-2\pi i \frac{\omega}{\alpha}(-|\alpha| t)} \ |\alpha| dt, & \alpha < 0 \end{cases}$$

$$= \begin{cases} \dfrac{1}{|\alpha|} \displaystyle\int_{y=-\infty}^{+\infty} f(y) e^{-2\pi i \frac{\omega}{\alpha} y} \ dy, & \alpha > 0 \\ \dfrac{1}{|\alpha|} \displaystyle\int_{y=-\infty}^{+\infty} f(y) e^{-2\pi i \frac{\omega}{\alpha} y} \ dy, & \alpha < 0 \end{cases} = \frac{1}{|\alpha|} F\left(\frac{\omega}{\alpha}\right).$$

Тиме је тврђење доказано. ■

Следећа је „Рејлејева теорема“ о енергији. Она каже да је интеграл квадрата величине функције (тј. енергија функције) једнак интегралу квадрата величине њене Фуријеове трансформације.

Теорема 5. (Рејлијева) $\int_{-\infty}^{+\infty} |f(t)|^2 \, dt = \int_{-\infty}^{+\infty} |F(\omega)|^2 \, d\omega$.

Доказ:

$$\int_{-\infty}^{+\infty} |f(t)|^2 \, dt = \int_{-\infty}^{+\infty} f^*(t) f(t) \, dt =$$

$$= \int_{t=-\infty}^{+\infty} \left[\int_{\omega=-\infty}^{+\infty} F(z) e^{2\pi i \omega t} \, d\omega \right]^* f(t) \, dt = \int_{t=-\infty}^{+\infty} \left[\int_{\omega=-\infty}^{+\infty} F^*(\omega) e^{-2\pi i \omega t} \, d\omega \right] f(t) \, dt$$

$$= \int_{\omega=-\infty}^{+\infty} F^*(\omega) \left[\int_{t=-\infty}^{+\infty} e^{-2\pi i \omega t} f(t) \, dt \right] d\omega = \int_{z=-\infty}^{+\infty} F^*(\omega) F(\omega) \, d\omega = \int_{-\infty}^{+\infty} |F(\omega)|^2 \, d\omega.$$

Дакле, $\int_{-\infty}^{+\infty} |f(t)|^2 \, dt = \int_{-\infty}^{+\infty} |F(\omega)|^2 \, d\omega$, што је и требало доказати. ■

Рејлејева теорема о енергији зове се и Парсевалова теорема за енергетске сигнале. Она вели да је снага сигнала једнака збиру квадрата величина различитих хармонијских компоненти присутних у (дискретном) спектру. Та нам, 5. теорема, говори да снагу сигнала одређују његови коефицијенати Фуријеовог реда, или другим речима, хармоници присутни у сигналу.

Следеће је својство Фуријеове трансформације да извод у временском домену дате функције види као множење Фуријеове трансформације те функције фактором $2\pi i$ у фреквенцијском домену.

Теорема 6. (Диференцијације) $\mathcal{F}\{f'(t)\}(\omega) = 2\pi i \omega F(\omega)$.

Доказ:

$$f(t) = \mathcal{F}^{-1}\{F\} = \int_{\omega=-\infty}^{+\infty} F(\omega) e^{2\pi i \omega t}\, d\omega$$

$$f'(t) = \int_{\omega=-\infty}^{+\infty} 2\pi i \omega F(\omega) e^{2\pi i \omega t}\, d\omega$$

$$f' = \mathcal{F}^{-1}\{2\pi i \omega F(\omega)\}$$

$$\mathcal{F}\{f'\} = 2\pi i \omega F(\omega)$$

а то је и требало доказати. ■

Знамо да непрекидну Фуријеову трансформација (3. дефиниција) можемо разумети као прелазак са временског t на фреквенцијски домен ω, при чему копија не губи информације оригинала. Уз то приметимо да множење комплексном јединицом $e^{i\omega t}$ временских вредности $f(t)$ врши смицање, ротацију ових вредности у имагинарној равни које тек добија „физички смисао“, тачније смисао у (мојој) „теорији информације“ (The six dimensions of space-time). Теорема диференцијације томе придодаје (до даљњег рецимо квази) кинетичку енергију (Angular frequency).

Друга перспектива коју нам ове трансформације отварају су фреквенције. Оне су на (засад псеудо) начин присутне и на рецимо трајекторијама где их иначе не виђамо (Frequency). Трећи концепт су канали преноса, које представљају овакве трансформације.

50. Мелинова трансформација

За разлику од Фуријеових и Лапласових трансформација које су уведене за решавање физичких проблема, Мелинова трансформација је настала у математичком контексту. Роберт Мелин био је фински математичар који је студирао код Карла Вајерштраса.

Дефиниција 1. Нека је $f(t)$ функција дефинисана на позитивној реалној оси, $0 < t < \infty$. Мелинова трансформација $\mathcal{M}$ је пресликавање функције f у функцију F дефинисано у комплексној равни релацијом

$$\mathcal{M}\{f\}(s) = F(s) = \int_0^\infty f(t) t^{s-1}\, dt$$

Функција $F(s)$ је Мелинова трансформација функције $f(t)$. Обрнуто, када је дато $F(s)$, онда је

$$\mathcal{M}^{-1}\{F\} = f(t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} t^{-s} F(s)\, ds$$

Инверзна Мелинова трасформација. ▨

Ознака другог интеграла имплицира да је он линијски интеграл узет преко вертикалне линије у комплексној равни, чији прави део c треба да задовољи само благу доњу границу. Услови под којима ова инверзија важи дати су у Мелиновој теореми о инверзији.

Основни, горњи интеграл након смене варијабле $t = e^{-x}$, одакле $dt = -e^{-x} dx$ и $f(e^{-x}) = g(x)$ постаје двострана Лапласова трансформација

$$\mathcal{L}\{g\} = \int_{-\infty}^{\infty} g(x) e^{-sx}\, dx$$

што пишемо

$$\mathcal{M}\{f(t); s\} = \mathcal{L}\{f(e^{-x}); s\}.$$

Да би добили Фуријеову трансформацију, у ову смену ставимо $s = \alpha + 2\pi i\beta$

$$F(s) = \int_{-\infty}^{\infty} f(e^{-x}) e^{-\alpha x} e^{-2\pi i\beta x}\, dx$$

Резултат је

$$\mathcal{M}\{f(t); \alpha + 2\pi i\beta\} = \mathcal{F}\{f(e^{-x}) e^{-\alpha x}; \beta\}$$

Иначе је Фуријеова трансформација

$$\mathcal{F}\{f; \beta\} = \int_{-\infty}^{\infty} f(x) e^{-2\pi i\beta x}\, dx.$$

Дакле, за дату вредност $\mathrm{Re}(s) = \alpha$ која припада траци дефиниције, Мелинова трансформација функције се може изразити и као Фуријеова трансформација. Оно на шта из овога треба обратити пажњу јесу пресликавања амплитуде, тада вероватноће Лапласове и Фуријеове трансформације у информацију Мелинове.

Наиме, горњи преписи Мелинове у двострану Лапласову трансформацију су:

$$\mathcal{L}\{f\}(s) = \mathcal{M}\{f(-\log x)\}(s), \quad \mathcal{M}\{f\}(s) = \mathcal{L}\{f(e^{-x})\}(s);$$

а када се у ове преписе укључе и Фуријеове трансформације:

$$\begin{cases} \mathcal{F}\{f\}(-s) = \mathcal{L}\{f\}(-is) = \mathcal{M}\{f(-\log x)\}(-is) \\ \mathcal{M}\{f\}(s) = \mathcal{L}\{f(e^{-x})\}(s) = \mathcal{F}\{f(e^{-x})\}(-is) \end{cases}$$

а то је само још једно од већ постојећих многобројних тумачења ових једнакости.

Међутим, аутентнични примери Мелинове трансформације многобројни су и неизбежни у њеној каснијој допуни кроз „теорију информације“.

Пример 1. (Cahen–Mellin integral) Мелинова трансформација функције $f(x) = e^{-x}$ је гама функција $\Gamma(s) = \int_0^\infty x^{s-1} e^{-x}\, dx$. Она је дефинисана за све комплексне бројеве, а $\Gamma(n) = n!$ дакле факторијел ($n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$) за позитивне целе бројеве. ❑

Пример 2. (Полином) За позитивне експоненте ($\lambda > 0$) интеграл $\int_0^\infty x^\lambda\, dx$ није конвергентан, али се пригодна функција може дефинисати са

$$f(x) = \begin{cases} x^\lambda, & x < 1 \\ 0, & x > 1 \end{cases}$$

па су њене Мелинове трансформације:

$$\mathcal{M}\{f(x)\}(s) = \begin{cases} \int_0^1 x^{s-1} x^a\, dx = \frac{1}{s+a}, & \mathrm{Re}(s) > -a \\ \int_1^\infty x^{s-1} x^b\, dx = -\frac{1}{s+b}, & \mathrm{Re}(s) < -b \end{cases}$$

Проверите! ❑

51. Интегралне трансформације

Интегрална трансформација је пресликавање који производи нову функцију $g(y)$ интеграцијом производа кернел функције $k(x,y)$ и постојеће функције $f(x)$ у датим границама. Процес назван трансформацијом дат је једначином

$$g(y) = \int_a^b k(x,y) f(x)\, dx$$

Неке од тих трансформација поменуте су у недавним насловима. Једна је Лапласова са кернелом e^{-xy} и границама интеграције од нуле до плус бесконачно, друга Фуријеова кернела $e^{-ixy}/\sqrt{2\pi}$ са границом од минус до плус бесконачно, а трећа Мелинова која их поопштава.

Са становишта преноса информација интегралне трансформације се могу једнако посматрати на неколико битно различитих начина. Један је овде наглашаван [43. Условне вероватноће], где смо кернел третирали као условну вероватноћу $k(x,y) = r(y|x)$, а постојећу функцију попут густине вероватноће $f(x) = \rho(x)$. У наставку идемо том интерпретацијом интегралних трансформација, као канала у ужем смислу, али приметићемо да тако дефинисан канал има једну непрецизност и одмах ћемо је расправити. Са променом улаза мења се дефиниција канала, а ми даље раздвајамо ту улазну информацију од канала у ужем смислу.

Укратко, вероватноћа да y падне у интервал $(y, y+dy)$ ако је x сигурно било у $(x, x+dx)$ износи $r(y|x)dy$. Условна расподела $r(y|x)$ не зависи од улазне расподеле x, па згодно дефинише канал, пренос $x \to y$, односно канал у ужем смислу. Ако је $\rho(x)dx$ вероватноћа да x буде у $(x, x+dx)$ то је $r(y|x)\rho(x)dxdy$ вероватноћа да y буде у интервалу $(y, y+dy)$. Према томе

$$\varphi(y)dy = r(y|x)\rho(x)dxdy$$

односно

$$\varphi(y) = \int_{-\infty}^{+\infty} r(y|x)\rho(x)\, dx$$

за произвољно y. Тако одређујемо густину излаза (φ) из канала (r) кроз који шаљемо произвољну расподелу (ρ) улазних сигнала.

За разлику од претходне дефиниције, канала у ширем смислу [43.], сада исти канал дефинишемо у ужем смислу и то аналогно третирању дискретних канала. Препознајемо Фредхолмову интегралну једначину прве врсте, ознаке $k(x,y) = r(y|x)$, где је

$$\int_{-\infty}^{+\infty} k(x,y)\, dy = 1, \quad k(x,y) \ge 0.$$

Ако је произвољна функција $k(x,y)$ интеграбилна по оба аргумента (x и y) и нормирана по другом аргументу (y), тада она може дефинисати канал у ужем смислу. Због услова нормираности улазне расподеле (ρ) видимо да је функција φ, ако постоји, такође нормирана.

Пример 1. Када је канал дефинисан функцијом која не зависи од друге променљиве, на излазу је константна расподела, она не реагује на промене вредности улазне расподеле.

Наиме, ако је $k(x,y) = f(x) \geq 0$ за све $x \in (a,b)$ и $k(x,y) = 0$ за све остале $x \notin (a,b)$, чак и ако f није нормирано на јединицу, добијамо излазну расподелу:

$$\varphi(y) = \int_{-\infty}^{+\infty} k(x,y)\rho(x)\,dx = \int_a^b f(x)\rho(x)\,dx = \text{const},$$

дакле константу за сваку улазну расподелу ρ. Вредност константе одређује нормираност функција ρ и φ. Добијамо униформну расподелу на интервалу (a,b) ако је улазна расподела ненегативна на истом интервалу. ❑

Пример 2. Када је канал дефинисан функцијом која не зависи од прве променљиве, на излазу је расподела самог канала, без обзира на улазну расподелу.

Наиме, ако је $k(x,y) = f(y)$, за произвољну улазну густину ρ, због њене нормираности, добијамо излазну расподелу $\varphi(y) = f(y)$. ❑

У оба примера имамо бескорисне канале (црне кутије), јер на основу примљеног излаза не да се одредити карактер улазних података. Први од примера, парафразирам, каже да без „апарата“ за пренос информације нема информације, а други да без неког таквог нема нити дезинформације. Специјални случај претходних је канал $k(x,y) = \text{const}$, а који производи униформну расподелу на некој области Oxy равни.

Кернел $k(x,y)$ дефинише линеарни оператор $\mathrm{K}: \rho \to \varphi$ који скуп улазних расподела пресликава у скуп излазних расподела. Наиме:

$$\mathrm{K}\{a\rho_1 + b\rho_2\} = \int_{-\infty}^{+\infty} k(x,y)[a\rho_1(x) + b\rho_2(x)]\,dx =$$

$$= a\int_{-\infty}^{+\infty} k(x,y)\,\rho_1(x)\,dx + b\int_{-\infty}^{+\infty} k(x,y)\,\rho_1(x)\,dx = a\mathrm{K}\{\rho_1\} + b\mathrm{K}\{\rho_2\}.$$

То је општа особина која, наравно, важи и за Лапласову [48.] или Фуријеову [49.] трансформацију.

На пример, када је канал у ужем смислу дефинисан експоненцијалном расподелом

$$k(x,y) = xe^{-xy}$$

са позитивним варијаблама ($x, y > 0$), тада се K може свести на

$$\mathcal{L}\{f(t)\} = F(s) = \int_0^\infty f(t)e^{-st}\,dt$$

дакле на Лапласов оператор (трансформацију).

Пример 3. Када је улазна расподела Ерлангова ($\lambda, x > 0$ са $m = 0,1,2,\ldots$)

$$\rho_m(x) = \frac{\lambda(\lambda x)^m}{m!} e^{-mx}$$

полазећи од познате (табличне) једнакости за Лапласов оператор

$$\mathcal{L}\left\{\frac{t^n}{n!} e^{\alpha t}\right\} = \frac{1}{(s-\alpha)^{n+1}}$$

добијамо излазну расподелу

$$\varphi_m(y) = \frac{1}{(y+\lambda)^2}$$

са $y > 0$. ❑

Посебно, за $m = 0$, улазна Ерлангова расподела своди се на експоненцијалну [40.], а излазна постаје Кошијева расподела

$$\varphi(y) = \frac{\lambda}{(y+\lambda)^2}$$

са $y > 0$. Лако је проверити да је Кошијева расподела дивергентна, тј. нема коначне моменте (очекивање, дисперзију, итд.), па према [41. Теорема 1.] не мора имати коначну информацију континуума.

Познато је да време услуживања клијената, време услуге купца, време поправке аутомобила, време разговора телефоном — има експоненцијалну или Ерлангову расподелу. Канал у ужем смислу са експоненцијалном расподелом се, према томе, имитира трговца чије је појединачно време услуживања y, а просечно време услуге $1/x$. Излазна расподела је средња вредност (ρ) функције канала у ужем смислу.

Канал са нормалном (Гаусовом) расподелом

$$k(x,y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-y)^2}{2\sigma^2}}$$

са $x, y \in \mathbb{R}$, има нешто другачије интерпретације. То је расподела која добро описује грешке мерења или гађања. Ако меримо величину чија је тачна вредност x, а средња грешка мерног уређаја износи σ, онда вероватноћа да ће бити измерен број у интервалу $(y, y+dy)$ износи $k(x,y)dx$.

Сам канал K, који год, садржи или производи неодређеност чија је средња вредност

$$\mu(\mathrm{K}) = -\int_{-\infty}^{+\infty} k(x,y) \log k(x,y)\, dy,$$

а када је канал нормалне расподеле имамо [36. Расипање] његову информацију

$$S(\mathrm{K}) = \log \sigma\sqrt{2\pi e}$$

заправо једнаку шуму који он производи. То је и резултат једнак [40. Пример 1.], а може се добити на још један начин.

Наиме, средња вредност неодређености канала по улазној расподели [43.] је шум:

$$S(X,Y) = S(Y|X) = \int_{-\infty}^{+\infty} \rho(x) S(\mathrm{K})\, dx = S(\mathrm{K})$$

где се подразумева непрекидни канал, па нема потребе да користимо ознаку S^* као пре. Овде је $S(\mathrm{K})$ константа, а ρ је нормирано. Шум нормалног канала не зависи од улазне расподеле.

Затим, произвољна излазна расподела, варијансе σ_y^2, према [41. Теорема 1.] има информацију

$$S(Y) \le \log \sigma_y \sqrt{2\pi e}$$

а са друге стране, она је једнака збиру средње узајамне информације и шума, те је

$$0 \le S(X,Y) \le \log \frac{\sigma_y}{\sigma}.$$

Добијене релације важе за сваку улазну и излазну расподелу кроз нормални канал, а последња (десна) неједнакост постаје једнакост ако је излазна расподела нормална. Посебно отуда видимо да је $\sigma_y/\sigma \ge 1$ те је за нормалну излазну расподелу

$$S(\mathrm{K}) \le S(Y).$$

Интуитивно очекујемо такође да ће неодређеност, па тиме и њихова реализација у информацију бити већа када се на основну иноформацију дода шум канала, а ови детаљи са формулама ту су да укажу на склад теорије.

Да бисмо упоредили улазну и излазну информацију подсетимо се да је за нормалне расподеле

$$\int_{-\infty}^{+\infty} k(x,y;\sigma_1^2) k(x,z;\sigma_2^2)\, dy = \frac{1}{\sigma_1 \sigma_2 2\pi} \int_{-\infty}^{+\infty} e^{-\frac{(x-y)^2}{2\sigma_1^2}} e^{-\frac{(y-z)^2}{2\sigma_2^2}}\, dy = \dots \to k(x,-z;\sigma_1^2+\sigma_2^2)$$

За то постоје разни докази. Дакле, серијски спој нормалних канала је нормални канал, па ако је у таквом каналу улазна расподела нормална, такође је и излазна нормална. Прецизније, приметимо из $\sigma_y^2 = \sigma_x^2 + \sigma^2$, излазна варијанса (дисперзија) збир је улазне и варијансе самог канала, из чега произилази поредак информација

$$S(X) \le S(Y),$$

што је још једна потврда претходно реченог.

Контрапримером се може показати да последње неједнакости не важе за не-нормалне улазне расподеле у нормални канал. Ако је улазна расподела униформна на датом интервалу, тада излазна расподела није униформна, јер би то противречило [41. Теорема 1.].

V Део: Композиције

Питања попут следећа два инспирација су ми да тему „сила неодређености“ разрађујем на начине наставка текста, који са претходним открићем такође добијају дах новине.

Питање: Могућности су више од простих опција – кажете, можете ли ми то појаснити?

Одговор: Могућности нису „само“ опције, оне су и (конструктивна или деструктивна) снага „силе неизвесности“. Отуда слободе које имамо (као и живих бића уопште) оснажују нас да пркосимо спонтаности, иначе мртвих тела да се препуштају „као кладић низ воду“. Слободе су заиста потенцијал народа (живог света) налик оном гравитационе, или Кулонове силе.

У прилогу (Uncertainty Force) су докази. Прво, наводне силе неизвесности опонашају Кеплеров други закон (потег сунце-планета у једнаким временима пребрише једнаке површине). Друго, у свом „простору вероватноћа“ оне узрокују кретања по коникама (хиперболе, параболе, елипсе, укључујући кружнице и праве линије – као у физичком свету сила). Треће, дејство неизвесности шири се брзином светлости и опада са квадратом удаљености.

Интуитивно, то исто видимо из „спонтаног“ чешћег дешавања вероватнијих исхода. Пишем наводног, јер овим открићем она (спонтаност) постаје врста силе, а само откриће „чудно“ и „одбојно“, бар на први поглед, као и својевремено откриће гравитације. Са овим добија се на значају и смислу бјективности самих опција, некаузалности овог света, а тиме и на „паралелним реалностима“.

Питање: О каквој „брзини светлости“ ширења дејства „силе неизвесности“ говорите у апстрактном „простору вероватноћа“; зар светлост и брзина нису појмови физичке реалности?

Одговор: Прочитајте пажљиво те прилоге. Апстракција је рецимо „виртуелна сфера“ [38.]. Ма којег полупречника, она дели „простор вероватноћа“ на веће унутар и мање вани вероватноће, математички је појам, статичан, али је у (произвољној датој) примени везана за посматрача. Као објекат без сопственог времена (попут светлости), она путује брзином садашњости самог субјекта, посматрача. А та је брзина светлости.

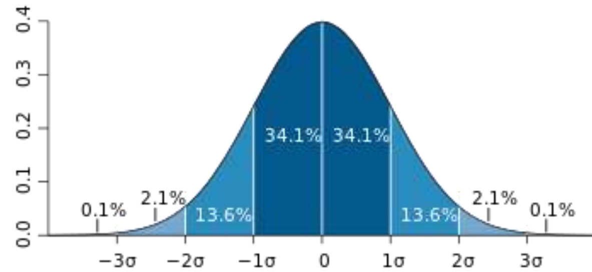
Брзином светлости садашњост, ма ког посматрача из физичке реалности, путује из прошлости ка будућности. Доследно, појаве које описујемо теоријом вероватноће, а можемо их посматрати, шире своја дејства брзином светлости.

52. Дифузија

Познати ставови о збиру нормалних расподела случајних варијабли поучни су на разне начине. Једноставан случај ($\mu = 0$ и $\sigma = 1$) нормалне, тзв. Гаусове, расподеле дефинише густина

$$\rho(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

чији граф је приказан на слици десно.



Посебно за ову дистрибуцију, вредности мање од једне дисперзије (σ), тачније речено њене стандардне девијације, од средње вредности (μ), чине 68,27% укупних могућности. Док две стандардне девијације од средње вредности чине 95,45% читавог скупа, а три чине 99,73%.

Премештањем (транслацијом) дуж хоризонталне, x-осе, или развлачењем/скупљањем графа, које постижемо сменом $x \to (x - \mu)/\sigma$, поменути односи дисперзије и очекивања (σ и μ) не мењају се. Тада густина нормалне расподеле постаје

$$\rho(x; \mu, \sigma) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

расподеле $\mathcal{N}(\mu, \sigma^2)$ случајних варијабли X sa математичким очекивањем μ и девијацијом σ. Она је симетрична око вертикалне осе $x = \mu$.

Када имамо две случајне варијабле, X_1 и X_2, свака своје нормалне расподеле, редом $\mathcal{N}(\mu_1, \sigma_1^2)$ и $\mathcal{N}(\mu_2, \sigma_2^2)$, њихов збир је опет случајна варијабла $Z = X_1 + X_2$ нове и такође нормалне расподеле $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. Нема потребе да доказе преписујем овде, али поменућу да се статистика са хи-квадрат расподелама заснивају на овом ставу.

Основна хи-квадрат расподела се добија као збир квадрата N независних Гаусових (нормалних) случајних промењљивих са нултом средњом вредношћу, а генерализације хи-квадрат расподеле се изводе као и код Гаусове

$$\chi^2(X) = \sum_{k=1}^{N} \left(\frac{x_k - \mu_k}{\sigma_k}\right)^2$$

сабирањем квадрата других типова одговарајућих случајних промењљивих.

Приметимо још једну занимљивост Гаусове расподеле, да се исти став о повећању очекивања и дисперзије може доказивати помоћу конволуција

$$\rho_Z(z) = \int_{-\infty}^{+\infty} \rho_2(z - x)\rho_1(x)\, dx.$$

Те нормалне густине ρ_1 и ρ_2 својом „композицијом“ (збиром) дају густину ρ_Z конволуције, опет нормалне расподеле $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. Занимљивост је у запису конволуције и израчунавању,

који изгледају као да радимо са множењем нормалних расподела, стварном композицијом функција њихових густина.

Производ две Гаусове (нормалне) случајне променљиве има дистрибуцију линеарне комбинације две хи-квадрат случајне променљиве

$$X_1X_2 = \frac{1}{4}(X_1 + X_2)^2 - \frac{1}{4}(X_1 - X_2)^2.$$

Ту су X_1, X_2, $X_1 + X_2$ и $X_1 - X_2$ нормалне променљиве, али њихови квадрати имају хи-квадрат расподелу. Та „скоро сличност“ дисперзије χ-квадрат расподеле и дисперзије Гаусове расподеле долази од саме опште дефиниције дисперзије [26.], варијансе $\sigma^2(X) = \mathrm{Var}(X)$, која је средња вредност (очекивање) квадрата одступања од средње вредности μ саме расподеле. У дискретном случају

$$\mathrm{Var}(X) = \sum_{k=1}^{n} p_k(x_k - \mu)^2$$

где је $p_k = \mathrm{Pr}(X = x_k)$ вероватноћа да ће k-та променљива бити реализована.

Разлика између хи-квадрат расподеле и варијансе подсећа на разлику између коваријансе и корелације. Додајмо овоме да се та „скоро сличност“ може приметити за неке динамичке процесе „теорије информације“, попут овде разматраних виртуелних сфера [38.]. Надовезивањем Гаусових расподела расте дисперзија резултирајуће расподеле и „композиција“ постаје „канал преноса“ са шумом који је једнак прираштају информације.

Таква врста канала преноса разликује се на пример од Марковљевих ланаца чији прираштај ланца је константан, по већој слободи начина појаве шума. Зато их је могуће применити на јединичну површину растуће „виртуелне сфере“, којој амплитуде опадају а са њима и вероватноће размене информације са евентуалним другим набојем, што би био процес не-константног ланца Маркова. Велике су могућности развоја теорије у том правцу и о томе ћу касније детаљније.

53. Дисперзија фигуре

Посматрамо виртуелну сферу у ширењу, или било коју сличну границу све мањих шанси неког простора вероватноћа. Рећи да се густина вероватноће на њој „разблажује" формално је једнако наизглед сасвим супротном, посматрању „ницања нових могућности" у датом делу површине. Нешто слично заиста се и дешава са вакуумом из којег ће искакати догађаји када га посматрамо све детаљније, али не морате се везати за ту физичку слику, за сада радимо апстрактно.

Знамо да се површина фигуре[9] повећава са квадратом њене дужине. То можемо визуелизовати са одређеним константним бројем равномерних детаља, рецимо тачака, неке растуће површине која је сразмерна квадрату удаљености датог пара тачака. У складу са тиме, дисперзија σ растућег дела сфере расте са њеним полупречником r.

У случају промене дисперзије сличне Гаусовој [40. Пример 1], информација је

$$S(\sigma) = \ln\left(\sigma\sqrt{2\pi e}\right)$$

а њена промена по јединици промене полупречника, њен извод по r, износи:

$$\frac{dS}{dr} = \frac{\sqrt{2\pi e}}{\sigma}\frac{d\sigma}{dr} = \frac{\sqrt{2\pi e}}{r}.$$

Информација дела сфере опада попут потенцијала.

То је хомотетија, врста пропорционалног пресликавања. Исти резултат, о информацији која опада на начин потенцијала, можемо добијати у случајевима „сличних фигура", на пример када пренос информације каналом има константан шум у односу на делове тока, као у Марковљевим ланцима. Информација коју такав канал преноси опада са бројем „карика ланца".

Условљеност следеће теореме „погрешном" расподелом, због чега се повећава информација кроз канал [14. Примери 1-3], односно настаје шум (сметња) преноса, показује да губици информације каналом преноса долазе од „неусклађености" коју сам канал [08. Адаптација каналом] покушава „поправити".

Теорема 1. Дате су три расподеле $\vec{p} = (p_1, \dots, p_n)$, $\vec{q} = (q_1, \dots, q_n)$ и $\vec{r} = (r_1, \dots, r_n)$ са три матрице $\hat{A} = [\![a_{ij}]\!]$, $\hat{B} = [\![b_{ij}]\!]$ и $\hat{C} = [\![c_{ij}]\!]$ реда n такве да је $\hat{A}\hat{B} = \hat{C}$, $\vec{r} = \vec{q}\hat{B}$, $\vec{q} = \vec{p}\hat{A}$. Тада за информације важи $S(\vec{r}|\vec{p}) \ge S(\vec{r}|\vec{q})$, где једнакост вреди ако и само ако је $b_{lj} = c_{ij}$ за свако $j = 1, \dots, n$.

Доказ: Из $c_{ij} = \sum_{l=1}^{n} a_{il}b_{lj}$ и коваријантних множења, за i-ту врсту информације матрице $\hat{C}$ је:

$$S_i(\hat{A}\hat{B}) = -\sum_{j=1}^{n} c_{ij}\log c_{ij} = -\sum_{j=1}^{n}\sum_{l=1}^{n} a_{il}b_{lj}\log c_{ij} = -\sum_{l=1}^{n} a_{il}\sum_{j=1}^{n} b_{lj}\log c_{ij} \ge$$

$$\ge -\sum_{l=1}^{n} a_{il}\sum_{j=1}^{n} b_{lj}\log b_{lj} = \sum_{l=1}^{n} a_{il}S_l(\hat{B}).$$

[9] Фигура је у геометрији било какав непразан скуп тачака.

Неједнакост долази из „погрешне расподеле“ [14.], због чега неједнакост постаје једнакост ако и само ако $b_{lj} = c_{ij}$ за свако $j = 1, \dots, n$. То ће бити нпр. када је $\hat{B}$ и $\hat{C}$ састављена од једног истог коефицијента, тј. када су врсте матрице $\hat{B}$ једнаке.

Даље, множећи са p_i добијамо:

$$\sum_{i=1}^{n} p_i S_i(\hat{A}\hat{B}) \geq \sum_{l=1}^{n} \sum_{i=1}^{n} p_i a_{il} S_l(\hat{B}) = \sum_{l=1}^{n} q_l S_l(\hat{B})$$

тј. $S(\vec{r}|\vec{p}) \geq S(\vec{r}|\vec{q})$, што је требало доказати. ■

У аналогијама са површинама, рећи ћемо да растом површине канал оригиналну информацију настоји „разблажити“, односно у другом случају да у каналу шум настаје „ницањем нових порука“. Апстрактност разматрања дозвољава овакве излете, из једних у друге области математике, јер свака тачна теорија у складу је са сваком другом тачном теоријом.

Занимљив модел таквог канала је васиона у интерпретацији моје теорије информације. Докле сеже могућност опажања прошлости, универзум је канал преноса информације. Сметње (шум) настају процесима успут, током протеклих 13,7 милијарди година. Физикална интерпретација овога, моје (хипо)тезе одавно, је да се супстанца „топи“ у простор, чиме се ентропија супстанце повећава а њена укупна информација смањује на уштрб физичког простора.

Подсећам, закон одржања за информацију при томе важи, као и материје у случају хабања било каквог физичког канала преноса порука, али не важе закони ентропије (други термодинамике) за простор као за супстанцу.

Пренос информација каналом дешава се и очитавањем резултата експеримената током процеса мерења. Слично настаје током сваког разговора док наша информација иде ка саговорнику, или обрнуто. Надам се да ово више није спорно, обзиром на 1. теорему, да сметње дате информације настају вишковима који се придружују успут, заједно са изобличењима послате због удруживања са околином. Другим речима, ако би се ток информације потпуно изоловао од средине којом тече, онда шума не би било. Информација не може тек тако у ништа нестајати, или из ничега настајати.

У реалним ситуацијама увек постоје сметње, па макар то била ретка насумична искакања честица из вакуума. Оне не морају „трајати“ да би могле дату информацију мењати (трајање је већ и само њихово присуство у променама). Са друге стране, закон одржања није нарушен ако упаде и мене гледамо у довољно широј околини.

54. Стационарни канал

У науци, рачунарству и инжењерству, црна кутија је систем улаза, излаза, или је карактеристика преноса, без икаквог знања о унутрашњем раду. Његова имплементација је „непрозирна“ (црна). Црна кутија овде је канал који не преноси информацију уопште. Стационаран канал био би онај који на излазу не повећава нити смањује улазну информацију.

Према претходном ставу [53. Теорема 1], такав је канал $\hat{C} = \hat{A}\hat{B}$ ако $b_{jk} = c_{ik}$ за сваку k-ту колону матрица $\hat{B}$ и $\hat{C}$, и сваки j-ти редак матрице $\hat{B}$. То је следећа ситуација

$$\hat{B} = \hat{C} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_1 & x_2 & \cdots & x_n \\ & & \vdots & \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

где је $\vec{x} = (x_1, x_2, \dots, x_n)$ нека (произвољне) расподела. За такве матрице важи $\hat{C}^2 = \hat{C} \cdot \hat{C} = \hat{C}$ и тада кажемо да је матрица идепотентна. За ове, множењем проверавамо:

$$\vec{p}\hat{C} = \vec{p}(\hat{A}\hat{B}) = (\vec{p}\hat{A})\hat{B} = \vec{q}\hat{B} = \vec{x}.$$

Исто детаљније:

$$\vec{p}\hat{C} = \left[(p_1 \quad p_2 \quad \cdots \quad p_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \right] \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_1 & x_2 & \cdots & x_n \\ & \vdots & & \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

$$= (q_1 \quad q_2 \quad \cdots \quad q_n) \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_1 & x_2 & \cdots & x_n \\ & \vdots & & \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} = (x_1 \quad x_2 \quad \cdots \quad x_n) = \vec{x},$$

јер матрица једнаких редова, коваријантно множена, не мења вредност вектора. Наиме:

$$\sum_{k=1}^{n} (q_k x_j) = \left(\sum_{k=1}^{n} q_k \right) x_j = x_j$$

јер је збир у другој загради један. Нема ништа од улазне информације $\vec{p}$ на излазу, јер је $\vec{x}$ било шта. Таква композиција, канал је црна кутија.

Разматрајући непрекидна пресликавања [42. Пренос варијабле], у 1. примеру, видели смо да је црна кутија и канал представљен густином који се може исказати производом две 1-дим густине, $g(x, y) = f_1(x) \cdot f_2(y)$. У њему следећем наслову [43. Условне вероватноће], препознали смо црну кутију у случају нултог коефицијента корелације, када је условна густина вероватноће била увек једнака завршној, независно од претходне. Те одреднице и ове у складу су, црна кутија је канал коме излаз не зависи од улаза.

Расправљајући о условним вероватноћама [43.], у недостатку јединствених ознака, користили смо израз „узајамна информација“, наспрам информације појединачне расподеле која је „својствена информација“. Тако првом описашмо процесе сликовито приказане изразима:

$$S(\vec{p} \to \vec{q}) = S(\vec{q} \to \vec{p}) = S(\vec{p}, \vec{q}).$$

Узајамна овде такође је блиска „корелацији“ теорије вероватноће. То је износ „информације“ који улаз достави излазу, или обрнуто речено, коју излаз прими од улаза. Са том, релације:

$$S(\vec{p} \to \vec{q}) + S(\vec{q}|\vec{p}) = S(\vec{q})$$

$$S(\vec{q} \to \vec{p}) + S(\vec{p}|\vec{q}) = S(\vec{p})$$

казују нам више о обиму својствене информације од $\vec{p}$ која је прошла из улаза до излаза канала. Сву осталу неодређеност, осим својствене почетне, која стиже као $\vec{q}$, називамо шумом.

Према томе, када имамо црну кутију, биће $S(\vec{q}|\vec{p}) = S(\vec{q})$, а онда из реченог $S(\vec{p}, \vec{q}) = 0$. Преноса порука између улаза и излаза (њихове комуникације) уопште нема. Таква је максимална сметња (чисти шум). Глувонеми канал (процес) представља матрица од једнаких колона (врста). Већ је то раније расправљено а овде углавном понављам.

Пре изношења следеће теореме, поновимом да хомоген ланац чини производ (композиција) једне те исте матрице, да је ланац Марковљев ако без обзира на то како је процес дошао у своје садашње стање он даје фиксна будућа стања (као код бацања новчића када следећи исход не зависи од претходних). Треће, матрица је потпуно стационарна ако се састоји од једнаких врста.

Став 1. Хомогени Марковљев ланац је потпуно стационаран ако и само ако је генерисан црном кутијом.

Доказ: Ако је за сваку расподелу $\vec{p}$ почетног корака $\hat{M}$, ланца $\vec{q} = \hat{p}\hat{M}\hat{M} \dots \hat{M}$, имамо константну резултирајућу $\vec{q}$, тада је $S(\vec{q}|\vec{p}) = S(\vec{q})$, тј. $S(\vec{p}, \vec{q}) = 0$, па је канал црна кутија и матрица M се састоји од истих колона. Обрнуто, ланац генерисан таквом матрицом је сасвим стационаран. ■

Када не бисмо знали како би изгледао ланац за различите почетне расподеле, онда ова теорема не вреди. Можемо бити сигурни једино да је стационарни низ $\vec{q}$ генерисан непознатом матрицом којој је он својствени вектор. Специјално је јединичној матрици својствени сваки вектор, па имамо другу крајност. Бела кутија генерише Марковљев ланац чији сваки корак има расподелу једнаку (произвољној) почетној.

У временским процесима, црна кутија је потпуно насумичан ток догађаја, када прошлост ни мало не одређује будућност, без памћења и без могућности предвиђања. Бела кутија је друга крајност, крајње детерминисаног тока догађаја. Сви остали процеси између су, имају бар мало случајности и бар мало извесности.

55. Коваријанса

Теореме нам могу много рећи, ако их знамо читати. Она из наслова „53. Дисперзије фигуре“ тако нам поручује да сметње у преносу поруке каналом не настају мањком него вишком информације. Вишком настаје мањак, жешћи потисну блаже. Шум је производ додат оригиналној информацији, осим код црне и беле кутије. Прва има максималну „дезинформацију“, а друга минималну.

Наравно да изражавање о „нестајању“ информације можемо користити и даље, али метафорички, као и појам „неистине“, или „празнине“ (вакуума). Такође, иста теорема са током доказа каже нам да квалитет преноса каналом зависи од адаптације канала на оно што се њиме преноси. Проста и прихватљива порука, коју иначе разумемо и без таквих теорема. Њена тачност или усклађеност са „очигледним“, међутим, има својих ограничења.

На пример, чињеница да и сам канал генерише неку информацију, означену са $S(\vec{p},\vec{q})$ у следећем наслову [54.], а коју не успевамо представити Шеноновом формулом и коефицијенатима матрице канала, говори о некој слабости. Претпостављам да се опис стварности ту показује као недовољан исечак извађен из контекста нечега чиме бисмо се даље радије бавили.

Тачна теорија неће бити у контрадикцији са било којом другом тачном теоријом, па сам ово место својевремено[10] разрешавао помоћу скаларног производа вектора одговарајућих опажаја субјекта и објекта, $S = a_1b_1 + a_2b_2 + \cdots + a_nb_n$, коју сам назвао „информацијом перцепције“. То је испало добро, сматрам, јер поменута „узајамна информација“ улазне и излазне поруке канала, пре личи на њу, на коваријансу или корелацију, него на Шенонову форму.

Коваријанса у теорији вероватноће и статистици представља меру јачине везе између промена две случајне променљиве, X и Y. Нека су очекивања (Expected value) ових средњих вредности, редом $E(X) = \mu_x$ и $E(Y) = \mu_y$, онда је њихова коваријанса

$$\mathrm{Cov}(X,Y) = E\big[(X-\mu_x)(Y-\mu_y)\big].$$

У дискретном случају, за скуп од $n = 2,3,4, \dots$ тачака података променљивих x и y, коваријанса се мери узимањем разлика између сваке од променљивих и њене средње вредности, које се затим множе у одговарајућим паровима и производи тих парова саберу а укупни збир подели са $n-1$. Сажето

$$\mathrm{Cov}(X,Y) = \frac{(x_1-\mu_x)(y_1-\mu_y) + (x_2-\mu_x)(y_2-\mu_y) + \cdots + (x_n-\mu_x)(y_n-\mu_y)}{n-1}.$$

Пример 1. Аналитичар компаније има податаке за пет квартала који показују раст бруто домаћег производа у процентима (X) и раст нове линије производа компаније у процентима (Y):

1. $x_1 = 2,\ \ y_1 = 10;$
2. $x_2 = 3,\ \ y_2 = 14;$
3. $x_3 = 2,\ \ y_3 = 13;$
4. $x_4 = 4,\ \ y_4 = 15;$
5. $x_5 = 5,\ \ y_5 = 16.$

[10] Р. Вуковић: „Информација Перцепције“, Економски институт Бања Лука, 2016.

Очекивања (Expectations) су средње вредности:

$$\begin{cases} E(X) = \mu_x = \dfrac{x_1 + x_2 + \cdots + x_n}{n} = \dfrac{2+3+2+4+5}{5} = \dfrac{16}{5} = 3{,}2 \\ E(Y) = \mu_y = \dfrac{y_1 + y_2 + \cdots + y_n}{n} = \dfrac{10+14+13+15+16}{5} = 13{,}6 \end{cases}$$

па је коваријанса података:

$$\mathrm{Cov}(X,Y) = \frac{(-1{,}2)(-3{,}6) + (-0{,}2)(0{,}4) + (-1{,}2)(-0{,}6) + (0{,}8)(1{,}4) + (1{,}8)(2{,}4)}{5-1} =$$

$$= \frac{(4{,}32) + (-0{,}08) + (0{,}48) + (0{,}72) + (4{,}32)}{4} = \frac{9{,}76}{4} = 2{,}44$$

Пошто је овде израчунао позитивну коваријансу, аналитичар може рећи да раст нове линије производа компаније има позитиван однос са кварталним растом БДП-а. ❑

Оба низа случајних променљивих $\vec{x} = (x_1, x_2, \dots, x_n)$ и $\vec{y} = (y_1, y_2, \dots, y_n)$ налазе се у неком n-дим (у горњем примеру димензије $n = 5$) апстрактном простору. Међутим, као вектори они разапињу паралелограм површине $|\vec{x} \times \vec{y}| = |\vec{x}||\vec{y}| \sin\varphi$, где је $\varphi = \sphericalangle(\vec{x}, \vec{y})$ угао између вектора. Разрађивао сам те геометријске начине својевремено, као у прилогу Varijance Vektor (2016), али и другачије.

Варијанса је посебан случај коваријансе када су две променљиве исте. Из прилога, али и из горњег примера види се да већој коваријанси одговара већа сагласност парова множених у сабирцима, а то значи да је већа ако множимо веће са већим разликама и мање са мањим, истог предзнака, па је она највећа ако множимо исте. У том максималном случају имамо варијансу

$$\mathrm{Var}(X) = E[(X-\mu)^2]$$

коју означавамо и са σ^2.

Када је коваријанса занемарљива (нула), онда се случајне варијабле узајамно понашају као да су у безакоњу. Сваки низ ради само за себе, насумично у односу на онај други. Коваријанса тада може представљати нулту „узајамну информацију“, $S(\vec{p}, \vec{q}) = 0$. Канал преноса информација, $X \to Y$, а и обрнуто, производи толико шума да од дезинформација почетна информација бледи. Варијанса нам говори о пропусности сваког канала са неком његовом сопственом максималном вредношћу и разабирање оригинала порастом шума слаби. У другој крајности, када дисонантност даље расте, коваријанса је негативна и мери супротну узрочност. Раст једне променљиве значи пад друге.

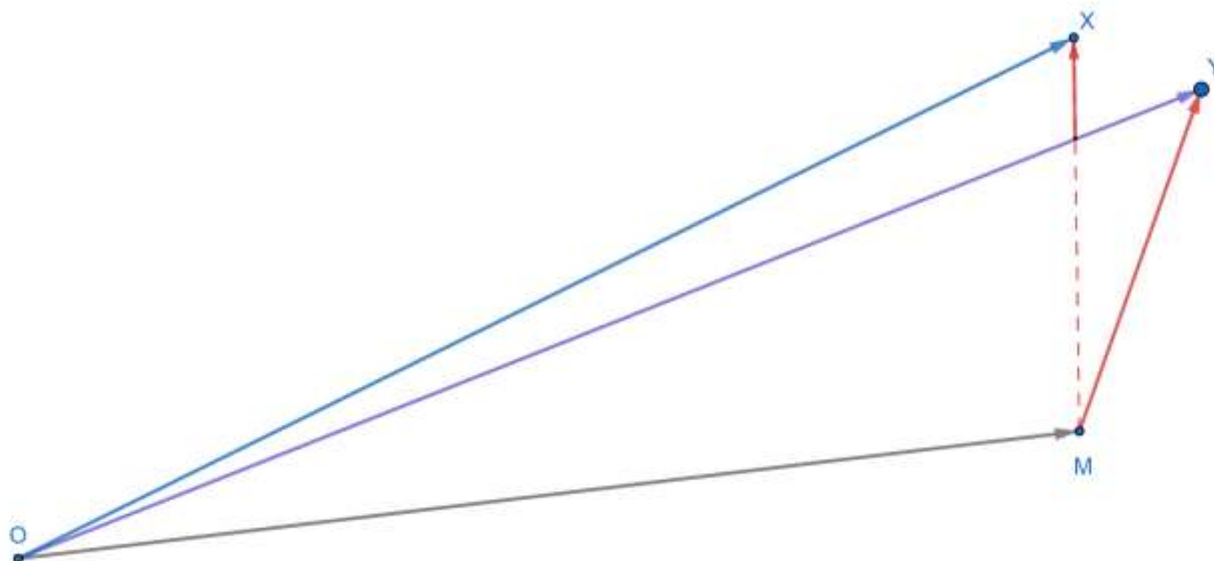
За још бољу симулацију узајамне информације послужиће корелација, са коефицијентом

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)\mathrm{Var}(Y)}.$$

Тај означавамо и са $\rho(X,Y)$. Предност корелације је, као што се види непосредно из формуле, јер је нормирана коваријанса. Увек је $-1 \le \rho(X,Y) \le 1$, а онда је лакше упоређивати дату корелацију са неком другом. Корак даље је информација перцепције.

56. Узајамна информација

У претходним насловима припремљено је објашњење узајамне информације помоћу коваријансе. На слици видимо n-дим прстор случајних тачака. То су низови, вектори $\overrightarrow{OX} = \vec{x} = (x_1, x_2, \dots, x_n)$ и $\overrightarrow{OY} = \vec{y} = (y_1, y_2, \dots, y_n)$ са заједничком средњом вредношћу, очекивањем $\mu = E(X) = E(Y)$ којом је дефинисан вектор $\overrightarrow{OM} = \vec{m} = (\mu, \mu, \dots, \mu)$.



Овим векторима дефинисани су вектори:

$$\begin{cases} \overrightarrow{MX} = \overrightarrow{OX} - \overrightarrow{OM} = (x_1 - \mu, x_2 - \mu, \dots, x_n - \mu) \\ \overrightarrow{MY} = \overrightarrow{OY} - \overrightarrow{OM} = (y_1 - \mu, y_2 - \mu, \dots, y_n - \mu) \end{cases}$$

који су окомити на вектор $\overrightarrow{OM}$. Ту сам окомитост, $\overrightarrow{OX}, \overrightarrow{OY} \perp \overrightarrow{OM}$, доказивао раније више пута, када тога није било около, рецимо под насловом „3.6 Стереометрија статистике“ књиге „Информација Перцепције“, па овде нећу. Битно је препознати да скаларни производ вектора разлика, $\overrightarrow{MX} \cdot \overrightarrow{MY}$, дефинише коваријансу, корелацију, затим узајамну информацију и информацију перцепције.

У случају да коефицијенти вектора $\vec{x}$ и $\vec{y}$ не припадају истој расподели вероватноћа, што је чешће у пракси, појавиће се два правоугла троугла OM_xX и OM_yY, са очекивањима:

$$\begin{cases} \mu_x = \dfrac{x_1 + x_2 + \dots + x_n}{n} \\ \mu_y = \dfrac{y_1 + y_2 + \dots + y_n}{n} \end{cases}$$

којима су дефинисане тачке M_x и M_y, односно вектори:

$$\overrightarrow{OM}_x = \vec{m}_x = (\mu_x, \mu_x, \dots, \mu_x), \quad \overrightarrow{OM}_y = \vec{m}_y = (\mu_y, \mu_y, \dots, \mu_y).$$

Ови вектори одузети редом од датих $\vec{x} = \overrightarrow{OX}$ и $\vec{y} = \overrightarrow{OY}$ дају векторе разлике:

$$\Delta\vec{x} = \vec{x} - \vec{m}_x, \quad \Delta\vec{y} = \vec{y} - \vec{m}_y$$

и постају вектори фактори коваријасе. Они се могу даље третирати са неким (сваким) заједничким почетком, назовимо га поново тачком M.

Било да су компактни као на горњој слици, или одвојени парови, скаларни производи вектора:

$$\overrightarrow{OX} \cdot \overrightarrow{OY} = |\overrightarrow{OX}||\overrightarrow{OY}| \cos \sphericalangle(XOY), \quad \overrightarrow{MX} \cdot \overrightarrow{MY} = |\overrightarrow{OX}||\overrightarrow{OY}| \cos \sphericalangle(XMY)$$

$$\vec{x} \cdot \vec{y} = |\vec{x}||\vec{y}| \cos \theta, \quad \Delta\vec{x} \cdot \Delta\vec{y} = |\Delta\vec{x}||\Delta\vec{y}| \cos \varphi$$

Ови еквивалентни парови записа који сличне интерпретације. Посебно, доле десно је узајамна информација $S(\vec{x}, \vec{y}) = \Delta\vec{x} \cdot \Delta\vec{y}$, скаларни производ вектора разлика који нам говори о односу улазне и излазне информације неког канала.

Када је број $S(\vec{x}, \vec{y})$ позитиван, тада су улаз и излаз сагласни, већи улазни бројеви даваће веће излазне и обрнуто мањи мање, а када је тај број негативан канал ради инверзију, већим улазним случајним вредностима додељује мање на излазу, а мањима веће. Шта се у каналу тачно дешава и како он то ради у тој интерпретацији је небитно.

Препознамо ли производ разлика, $\Delta\vec{x} \cdot \Delta\vec{y}$, даље као информацију перцепције, такву информацију можемо/морамо посматрати и као негативан број, а спискак мањкавости Шенонове дефиниције, ако је сматрамо неком општом, продужити. Са друге стране, нити узајамну информацију не треба третирати као општу, јер је она ограничена на различите појаве „истог“, на „улаз“ и „излаз“ њених вредности, односно стања пре и након процеса. Она се таква не односи на Шенонову изворно, средњу вредност дате расподеле вероватноћа.

Такође, примена производа разлика у информацији перцепције слично је ограничена. Стање разликујемо од процеса као што физички простор разликујемо од времена, или као што вектор (чија интерпретација је квантно стање) разликујемо од оператора (који интерпретира квантни процес) који га мења. Међутим, оператори (штавише, сви у математици) врста су вектора (дуални су онима на које делују), па поменуте посебности и нису суштинске.

Постоје дубље, суштинске, односно заједничке законитости које управљајуједанко како „обичним векторима“ тако и „операторима векторима“, стањима као и процесима, односно просторним као временским феноменима. На пример, једнако како по процесу можемо препознавати стања (не понашају се сви исто у истим условима), тако и стања бирају процесе (нпр. сила бира набоје).

Уосталом зато ми је било могуће на три просторне димензије доказати егзистенцију још три временске, или показати (Простор-Време) да би се једначине теорије релативности једнако понашале када бисмо из тог скупа шест просторно-временских димензија издвојили било које четири, од којих би три биле просторне а четврта временска имагинарна њима.

***

Напомињем да ова „узајамна информација“ није тачно она о којој сам писао 80-их и у књизи „Математичка теорија инфорамције и комуникације“ (немам је дигитално, потражите издање 1995. Удружења математичара Републике Српске у библиотеци), а која се и овде помиње као $S(X, Y)$, $S(XY)$, или $S(\vec{a} \to \vec{b})$, за разлику од нове $S(\vec{a}, \vec{b})$ где су $\vec{a}$ и $\vec{b}$ вектори са координататама

вероватноћама расподела случајних променљивих из скупова редом X и Y. У литератури, стара се обично означава са $H(X,Y)$, замало.

Тешко је у овом тренутку рећи којој би, $S(\vec{a},\vec{b})$ или $S(X,Y)$, назив „узајамна информација“ боље стојао. Ту ћемо дилему имати и касније, али за сада приметимо да машинско учење употребљава Пирсонову корелацију и да њој боље лежи $S(\vec{a},\vec{b})$. Међутим, та корелација је мера која проналази само линеарне односе и то сужава нашу примену у информатици, док не укључимо и статистике нелинеарних односа. Преме томе ће и разрада која следи [58.] бити сами почетак ове приче.

У наслову „43. Условне вероватноће“ уведен је класични појам узајамне информације $S(X,Y)$, са овде прихватљивима шумом канала $S(Y|X)$ и информацијом излаза $S(Y)$, и показано је да важе формуле попут

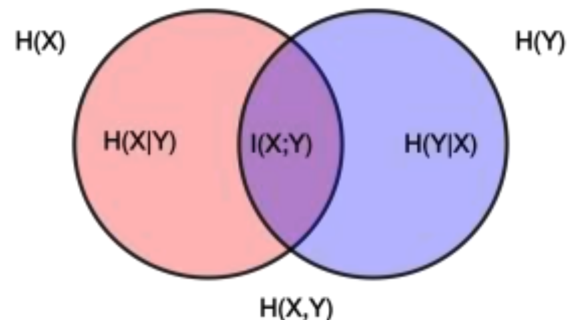
$$S(X,Y) + S(Y|X) = S(Y).$$

Ова како за непрекидне тако и дискретне информације. Аналогно, у литератури ћете најчешће (Mutual information) наћи формуле:

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X,Y) \\ &= H(X,Y) - H(X|Y) - H(Y|X) \end{aligned}$$

где су $H(X)$ и $H(Y)$ тзв. маргиналне ентропије (рецимо улазне и излазне), $H(X|Y)$ и $H(Y|X)$ су условне ентропије, а $H(X,Y)$ тзв. удружена ентропија (joint entropy).

На слици Венових дијаграма десно, помоћу скупова покушано је дочарати о чему се ради. Област оба круга је „удружена ентропија“, тј. $H(X,Y)$. Леви од кругова (црвен и љубичаст) је тзв. „индивидуална ентропија“ $H(X)$, рекли бисмо маргинална, или улазна. Црвени део била би „условна ентропија“ $H(X|Y)$. Круг на десној страни (плаво и љубичасто) је $H(Y)$, са плавим $H(Y|X)$. Љубичаста је та „међусобна информација“ $I(X;Y)$, у преводу и узајамна, заједничка, или обострана.



Лепо нацртана слика, осим што је нетачна. Ентропија и информација датог система нису заједно растуће величине, већ напротив. Да не дужим зашто, видите моје последње популарне прилоге о томе (http://rvukovic.net/blog/2202.html).

57. Опет Кеплер

У константним условима коефицијент корелације датих појава показиваће константне вредности, наравно, у оквиру статистички дозвољене грешке. Како корелацију, узајамну информацију, па на свој начин и информацију перцепције, генерише исти скаларни производ, $S(\vec{x},\vec{y}) = \Delta\vec{x}\cdot\Delta\vec{y}$, како смо то видели у претходном наслову [56.], то „константни услови" значе исте вероватноће датих догађаја, непромењен канал преноса информације, сталне перцепције. Сада ћемо ту законитост одржања, тако постављену, упоредити са „силом неизвесности".

Вероватноће се могу изводити директно из амплитуда, а тек секундарно везивати и проверавати на класичним вероватноћама математике. То у складу са квадратима модула амплитуда квантних стања, $\|\psi\|^2 = \psi^*\psi$, вероватноће налажења обзервабле[11] у квнтној механици. Сагласно томе су и „случајне варијабле" понекад „квадрати интензитета амплитуда". Објашњења самог тог Борновог закона давао сам у књизи Квантна механика.

Узимајући за случајне варијабле квадрате интензитета амплитуда, у општем случају комплексних бројева, очекивање се своди на:

$$E(\psi) = \mu_\psi = \frac{1}{n}\left(\|\psi_1\|^2 + \|\psi_2\|^2 + \dots + \|\psi_n\|^2\right) = \left\|\frac{\psi_1}{\sqrt{n}}\right\|^2 + \left\|\frac{\psi_2}{\sqrt{n}}\right\|^2 + \dots + \left\|\frac{\psi_n}{\sqrt{n}}\right\|^2$$

па ако пређемо на „скраћене векторе" ($\psi \to \psi' = \psi/\sqrt{n}$), биће $E(\psi) = nE(\psi')$. Пре евентуалног детаљнијег редефинисања класичне вероватноће на неку комплексну, зауставимо на ономе где смо већ стигли у претходним насловима. Коефицијент корелације је:

$$\rho(X,Y) = \mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)\mathrm{Var}(Y)} =$$

$$= \frac{\frac{1}{n}\sum_{k=1}^{n}(x_k - \mu_x)(y_k - \mu_y)}{\sqrt{\frac{1}{n}\sum_{k=1}^{n}(x_k - \mu_x)^2}\sqrt{\frac{1}{n}\sum_{k=1}^{n}(y_k - \mu_y)^2}} = \frac{\frac{1}{n}\Delta\vec{x}\cdot\Delta\vec{y}}{\frac{1}{n}|\Delta\vec{x}||\Delta\vec{y}|} = \frac{|\Delta\vec{x}||\Delta\vec{y}|\cos\varphi}{|\Delta\vec{x}||\Delta\vec{y}|} = \cos\varphi.$$

Коефицијент корелације је константан ако и само ако је угао $\varphi = \sphericalangle(XMY)$ између вектора $\Delta\vec{x}$ и $\Delta\vec{y}$ константан [56.]. Сличан резултат се добија сменом $\Delta\vec{x} \to \Delta\vec{x}' = \Delta\vec{x}/\sqrt{n}$, јер је:

$$\mathrm{Var}(X) = \frac{|\Delta\vec{x}|}{\sqrt{n}} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(x_k - \mu_x)^2} = \sqrt{\sum_{k=1}^{n}\left(\frac{x_k - \mu_x}{\sqrt{n}}\right)^2} = |\Delta\vec{x}'|.$$

Аналогно радимо са осталим случајним варијаблама ($Y \to Y'$). Израчунавамо и коваријансу[12]:

$$\mathrm{Cov}(X,Y) = \frac{\Delta\vec{x}}{\sqrt{n}}\cdot\frac{\Delta\vec{y}}{\sqrt{n}} = \frac{1}{n}\sum_{k=1}^{n}(x_k - \mu_x)(y_k - \mu_y) = \sum_{k=1}^{n}\left(\frac{x_k - \mu_x}{\sqrt{n}}\right)\left(\frac{y_k - \mu_y}{\sqrt{n}}\right) =$$

$$= \Delta\vec{x}'\cdot\Delta\vec{y}' = |\Delta\vec{x}'||\Delta\vec{y}'|\cos\varphi,$$

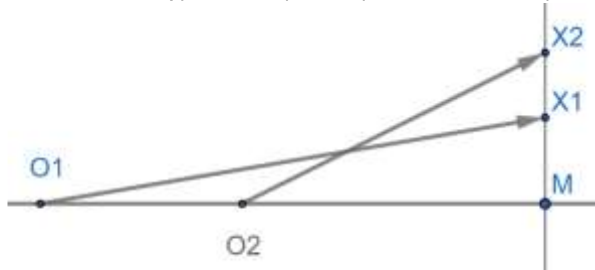
[11] Обзервабла – физички мерљива величина.
[12] Теорија вероватноће користи количник n уместо статистичког $n-1$.

јер угао између вектора (φ) не мења се самим скраћивањем вектора.

Крајњи прелазак на комплексну вероватноћу, онакве врсте која би одговарала квантној физици, није неопходан да приметимо горње резултате. Угао $\varphi = \sphericalangle(XMY)$ остаје константан, а аналогно следи са слике (претходног наслова). Текст непосредно наставља расправу о слици из претходног наслова, па главно не понављам. Све постаје јасније када су очекивања једнака и $\vec{m}_x = \vec{m}_y = \vec{m}$, затим је лакши прелаз на општији случај.

Растом броја n компоненти вектора $\vec{x} = (x_1, x_2, \dots, x_n)$ и $\vec{y} = (y_1, y_2, \dots, y_n)$ њихова очекивања (средње вредности) статистички се не мењају. Равномерно се скраћују вектори $\Delta\vec{x}'$ и $\Delta\vec{y}'$, правци им остају непромењени (према претходном) и угао који заклапају остаје исти. То значи да се дуж $\overline{XY}$, као и $\overline{X'Y'}$, из тачке M види под истим углом φ док се краци тог угла скраћују, сваки посебно сразмерно $\sqrt{n}$, па се (слични троуглови) скраћује и дужина $\overline{XY}$ том истом размером.

Даље посматрајмо веће троуглове, прво правоугли OMX. Док n расте, хипотенуза OX, као и катета OM, повећавају свака сразмерно $\sqrt{n}$, то се краћа катета смањује сразмерно коефицијенту $1/\sqrt{n}$.



По томе, синус оштрог угла у темену O смањује се сразмерно $1/n$ и површина троугла OMX не мења се. Површине троуглова O_1MX_1 и O_2MX_2 једнаке су, а то је управо оно што је овде већ доказивано генерисањем случајних бројева в. наслов „28. Центрифуга неизвесности“ и слику.

Да се веће странице троугла OMX, па тако и троугла OXY, повећавају сразмерно $\sqrt{n}$ (а не рецимо са n), видимо из следећих једнакости:

$$\overline{OM} = |\overrightarrow{OM}| = |\vec{m}| = |(\mu, \mu, \dots, \mu)| = \sqrt{\mu^2 + \mu^2 + \dots + \mu^2} = \mu\sqrt{1 + 1 + \dots + 1} = \mu\sqrt{n}$$

а очекивање (средња вредност) μ за једну исту расподелу не мења се без обзира на број исхода n.

Други већи троугао био је OXY. Његове странице OX и OY, са већим n све су дуже, а основица све краћа, па је угао θ, под којим се та основица XY види из темена O, све мањи. Опет се испоставља да је површина троугла OXY константна док се теме O удаљава од M, као на слици лево. То је као у случају слике десно наслова „24. Примери регресије II“, тамо доказивано софтвером а овде само геометријски.

У овој скрипти откриване и назване „силе неизвесности“ гурају случајне варијабле као набоје. Али заправо то су тачке у простору вероватноћа које се пресликавају, кажемо „померају“ по законима еквивалентим поопштеном Другом Кеплеровом (за једнака времена потег сила-набој пребрисаће једнаке површине), иначе важећи за све константне централне силе. Кулонове, или гравитационе, на пример.

58. Збир канала

Улазна информација у канал K је расподела вероватноћа $p_i = \Pr(\omega_i) \in (0,1)$, при чему је $\omega_i \in \Omega$ појединa порука индекса $i \in \{1,2,\dots,n\}$, дакле $\omega_i \in \{\omega_1, \omega_2, \dots, \omega_n\}$. Како се сабира информација $S(\vec{p})$ те расподеле $\vec{p} = (p_1, p_2, \dots, p_n)$ када је поделимо на тачно две здружене поруке вероватноћа $a = p_1 + \cdots + p_r$ и $b = p_{r+1} + \cdots + p_n$, прво је питање.

Теорема 1. Дата је расподела вероватноћа компонентама вектора $\vec{p} = (p_1, \dots, p_r, \dots, p_n)$ и поделом $a = p_1 + \cdots + p_r$, $b = p_{r+1} + \cdots + p_n$. Тада су $\vec{a} = (p_1/a, \dots, p_r/a)$, $\vec{b} = (p_{r+1}/b, \dots, p_n/b)$, $\vec{c} = (a, b)$ опет неке расподеле вероватноћа. При томе важи једнакост

$$S(\vec{p}) = aS(\vec{a}) + bS(\vec{b}) + S(\vec{c}),$$

где су функције S информације по Шенону.

Доказ: Да су $\vec{a}$, $\vec{b}$, $\vec{c}$, као и $\vec{p}$ расподеле вероватноћа очигледно је. Даље, користећи Шенонову дефиницију добијамо следеће три једнакости:

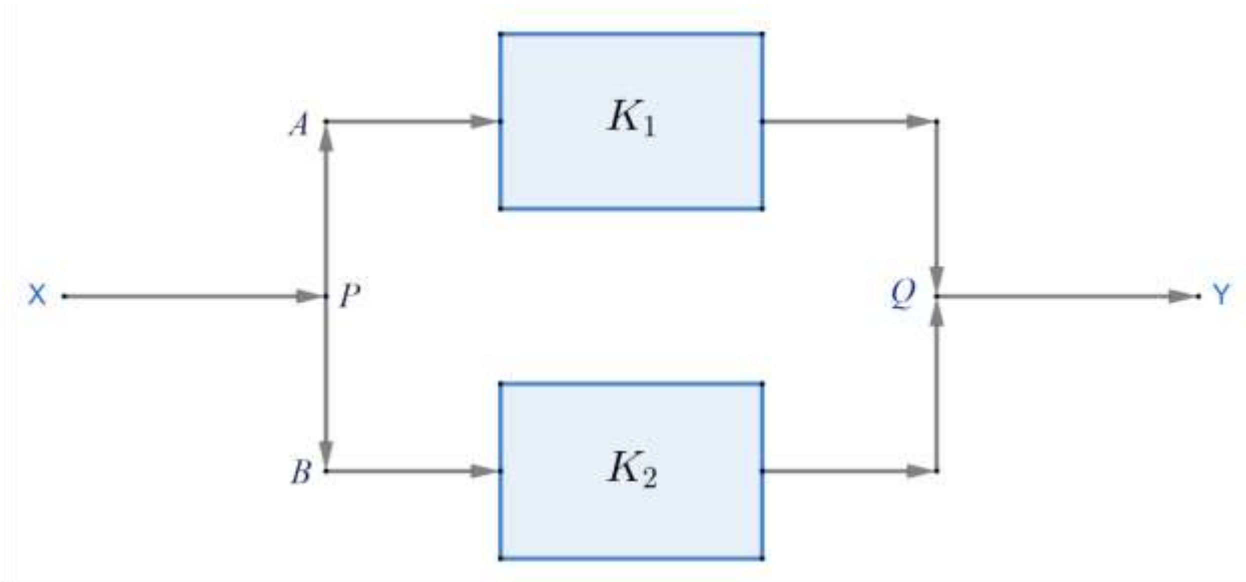
$$aS(\vec{a}) = -p_1 \log p_1 - \cdots - p_r \log p_r + a \log a$$

$$bS(\vec{b}) = -p_{r+1} \log p_{r+1} - \cdots - p_n \log p_n + b \log b$$

$$S(\vec{c}) = -a \log a - b \log b$$

из чијег сабирања следи $S(\vec{p}) = -p_1 \log p_1 - \cdots - p_n \log p_n$, што је требало доказати. ■

Применимо ово на канал који се састоји из два дела, $K = K_1 + K_2$, паралелне везе као на слици.



Сигнали долазе слева из X. У тачки P раздвајају се и део иде ка тачки A а остатак ка B. Део A улази у канал K_1 а B у канал K_2 након чега се излазни сигнали састају у Q и настављају десно ка Y. Према томе, канал $\hat{K} = [\![k_{ij}]\!]_n$ представља блок матрица

$$\hat{K} = \begin{pmatrix} \hat{K}_1 & 0 \\ 0 & \hat{K}_2 \end{pmatrix}$$

где су $\hat{K}_1 = [\![k_{ij}]\!]_r$ и $\hat{K}_2 = [\![k_{ij}]\!]_{n-r}$ дијагоналне квадратне матрице. Другим речима, $k_{ij} = 0$ ако је $i \le r$ и $j > r$, или ако је $i > r$ и $j \le r$. Улазни вектор је $\vec{p} = (p_1, \dots, p_n)$, излазни $\vec{q} = (q_1, \dots, q_n)$.

Вероватноће улазног вектора $\vec{p} = (p_1, \dots, p_r, \dots, p_n)$ канала K опет раздвајамо на збирове:

$$a = p_1 + \dots + p_r\,, \quad b = p_{r+1} + \dots + p_n$$

тако да имамо расподеле:

$$\vec{a} = \left(\frac{p_1}{a}, \dots, \frac{p_r}{a}\right), \quad \vec{b} = \left(\frac{p_{r+1}}{b}, \dots, \frac{p_n}{b}\right), \quad \vec{c} = (a, b)$$

као у горњој теореми и, према томе, улазна информација је $S(\vec{p}) = aS(\vec{a}) + bS(\vec{b}) + S(\vec{c})$.

Излазна расподела из канала $\hat{K} = [\![k_{ij}]\!]_n$ биће $\vec{q} = (q_1, \dots, q_r, \dots, q_n)$, такво да је $\vec{q} = \vec{p}\hat{K}$. Како је информација канала иначе

$$S(\hat{K}) = -\sum_{i,j=1}^{n} k_{ij} \log k_{ij}$$

то горња теорема важи и за излазне поруке, па можемо писати

$$S(\vec{q}) = a'S(\vec{a}') + b'S(\vec{b}') + S(\vec{c}')$$

са одговарајућим поделама датојЧ $a' = q_1 + \dots + q_r$, $b' = q_{r+1} + \dots + q_n$, $\vec{a}' = (q_1/a', \dots, q_r/a')$, $\vec{b}' = (q_{r+1}/b', \dots, q_n/b')$ и $\vec{c}' = (a', b')$, сада излазне расподеле.

Теорема 2. Шум збирног канала ($K = K_1 + K_2$) једнак је збиру шумова сабирака (K_1 и K_2).

Доказ: Информација i-те врсте канала је $S(\hat{K})_i = S_i = -\sum_{j=1}^{n} k_{ij} \log k_{ij}$ па за дату расподелу $\vec{p}$ средња вредност информација врста канала, што је и шум канала, је:

$$S(\vec{q}|\vec{p}) = \sum_{i=1}^{n} p_i S_i = a\sum_{i=1}^{r} \frac{p_i}{a} S_i + b \sum_{i=r+1}^{n} \frac{p_i}{b} S_i = S(\vec{a}'|\vec{a}) + S(\vec{b}'|\vec{b})$$

Дакле шум K једнак је збиру шумова канала K_1 и K_2. ■

Информацију збирног канала такође можемо сматрати једнаком збиру информација сабирака, јер улазна информација у каналу не нестаје тек тако, она се евентуално само дезинтегрише. Прилично је то и интуитивно јасно. Канал производи сметње, сопствене информације које такође не настају из ничега нити нестају у ништа. Укупна информација огригинала и укупни шум настао успут додају се. Међутим, информација је „количина опција“, а не рецимо нека „суштина“, скуп знања, па се може одржавати и тако што читаоцу постаје неразумљива.

Информација не тежи да буде „смисао“, или неки нама интелигентан облик података, сем када је принцип минимализма, или нека друга принуда тамо не доведе. Исто важи и за низ канала.

Надовезивањем исте поруке на нови исти канал настаје серијска веза (не паралелна као на слици). Њој одговара множење матрицом канала. Код простог дуплирања дате, имамо:

$$\hat{K} \cdot \hat{K} = \hat{K}^2 = \begin{pmatrix} \hat{K}_1 & 0 \\ 0 & \hat{K}_2 \end{pmatrix}^2 = \begin{pmatrix} \hat{K}_1^2 & 0 \\ 0 & \hat{K}_2^2 \end{pmatrix}$$

дакле опет блок матрицу, сада са квадратима првобитних блокова. Уопште, за дужи ланац је:

$$\hat{K}^n = \begin{pmatrix} \hat{K}_1 & 0 \\ 0 & \hat{K}_2 \end{pmatrix}^n = \begin{pmatrix} \hat{K}_1^n & 0 \\ 0 & \hat{K}_2^n \end{pmatrix}$$

што је лако доказати индукцијом. То је увек ланац са две серије, два режима рада. Информација послата серијом, иначе, па и кроз низ оваквих блок канала, прилагођава се базном каналу (K), она се адаптира и тежи својственој датог образца ових везаних канала. То је тежња шума. Бескрајним надовезивањем истог канала, канал се засићује додајући сопствену информацију укупној, стиже ка неком својственом максимуму и све више стагнира.

58.1. Капацитет канала

Максимална информација коју канал може пропуштати је капацитет канала. Капацитет канала, C, је класично (Шенон, 1948) описан као максимална брзина којом се информације могу преносити кроз канал, а те се одреднице можемо држати иако је њена првобитна намена била, а углавном и остала, сужена на брзине преноса симбола (слова) преко канала. У тој терминологији „ентропија“ случајне промењиве је просек обима „информација“, односно „изненађења“ или „неизвесности“ својствених могућим исходима варијабле.

1. Основна теорема бешумног канала гласи: „Ако језик има ентропију H (битова по симболу) и канал може преноси C бита у секунди, тада је могуће кодирати сигнал на начин да се преноси просечном брзином од $\frac{C}{H} - \varepsilon$ симбола у секунди, при чему $\varepsilon > 0$ можемо узимати произвољно малим. Пренос је немогућ просечном брзином већом од $\frac{C}{H}$.“

2. Основна теорема за дискретни канал са шумом: „Нека дискретни канал има капацитет C, а дискретни извор ентропију у секунди H. Ако је $H < C$, тада има начина кодирања тако да се излаз извора може пренети преко канала и уз произвољно малу фреквенцију грешака (или произвољно малу двосмисленост). Ако је $H > C$, произвољно мала вероватноћа грешке није достижна. “

Дакле, капацитет канала је максимална брзина информација коју канал може да пренесе. Мери се у битовима у секунди (bps) и грубо вреднује, процена узима у обзир само целу количину пренетих података, а изоставља квалитет комуникације.

Пропусни опсег (Bandwidth) се може сматрати подскупом термина капацитета канала. Када мери пропусни опсег, формула урачунава максималну количину информација која се може прецизно пренети у јединици времена. Међутим, на пример, капацитет канала може бити веома висок, али низак квалитет сигнала учиниће ниским пропусни опсег.

Шенон-Хартлијева теорема каже да је капацитет канала C теоријски најстрожа горња граница за брзину информација података која се могу пренети са произвољно малом грешком и просечном

снагом примљеног сигнала S кроз аналогни комуникациони канал подложан адитивном белом Гаусовом шуму (AWGN) снаге N

$$C = B\log_2\left(1+\frac{S}{N}\right)$$

где:

- C је капацитет канала у битима по секунди;
- B је пропусни опсег канала (bandwidth) у хецима;
- S је просечна снага примљеног сигнала преко пропусног опсега, мерена у ватима;
- N је просечна снага сметњи у пропусном опсегу, мерена у ватима;
- S/N је однос сигнал-шум, или носилац-шум на пријемнику.

На пример[13], за типичну телефонску линију (модем) са односом сигнал-шум од 30 dB и пропусним опсегом звука од 3 kHz, максимална брзина података је $C = 3000\log_2 1001 \approx 29{,}901$ kbps.

За сателитски ТВ канал са односом сигнал-шум од 20 dB пропусним опсегом видеа од 10 kHz, добијамо максималну брзину преноса података $C = 10\,000\,000\log_2 101 \approx 66{,}582$ kbps.

***

Мало другачији приступ је разумети капацитет канала као максимума узајамне информације

$$C = \max_{p(x)} I(X;Y)$$

где се максимум тражи у свим могућим улазним дистрибуцијама $p(x)$.

На пример, писаћа машина са сметњама и вероватноћом скретања 0,5 и 30 улазних симбола (можемо ићи са сваким другим словом и проћи са 15 бита). Израчунавамо:

$$C = \max_{p(x)} I(X;Y) = \max_{p(x)}[H(Y) - H(Y|X)] = \max_{p(x)} H(Y) - 1 = \log_2 30 - 1 = \log_2 15.$$

Уопште, за вероватноћу p, налазимо:

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - \sum_x p(x)H(Y|X=x) =$$

$$= H(Y) - \sum_x p(x)H(p) = H(Y) - H(p) \leq 1 - H(p)$$

дакле, $C = 1 - H(p)$.

Други пример, канал са шумом и не-преклапајућим излазима. $C = 1$ bit са расподелом $\vec{p} = \left(\frac{1}{2},\frac{1}{2}\right)$.

Такође, канал без шума и грешака двосмислености. $C = 1$ bit са расподелом $\vec{p} = \left(\frac{1}{2},\frac{1}{2}\right)$.

[13] L L Peterson and B S Davie, Computer Networks:a systems approach (Morgan Kaufmann), 1996. ISBN: 1-55860-368-9 (Paperback ISBN: 1-55860-404-9) pp 94-95.

***

Са становишта моје „теорије информације“ (незваничне) класични приступ капацитету канала је сувише ограничен на област (слова и симбола) где се углавном до данас и користи. Та класична теорија недовољно је примењива у области космологије, гравитације, електромагнетизма и сл, нити је на нивоу у доменима квантне физике, као што није нити у широј хемији, биологији, науци психе или заједница. Ни не тврди се да јесте, знам, али како то да примена информатике стиже са вештим инжењерима и програмерима и тамо где та теорија не добацује.

На пример, надмашивања класичног третирања капацитета канала (Beating the channel capacity) примећена су бар две деценије у квантној физици, али одговарајуће „теорије информације“ да те појаве објасни нема. Нема је, јер се из постојеће она не да извести.

Скок који нам је потребан је посматрање информације као ширег феномена, који се не може извести из њене досадашње скучености. Информација коју нудим је основно ткиво природе око нас (простора, времена, материје), а објективна неизвесност је њена суштина. Нуди се признање квантној механици да у расподели вероватноћа, може то бити и $\vec{p} = (p_1, \dots, p_n)$, њене компоненте треба даље разлагати, попут $p_k = \|\psi\|^2$, при чему величине ψ могу бити и комплексни бројеви (и не само они), а норме $\|\cdot\|$ не само еуклидске.

У том контексту, капацитет канала можемо видети кроз пресликавање $\vec{y} = \hat{A}\vec{x}$ тако да је

$$C = \max_x \|\hat{A}\vec{x}\|$$

где је улазни вектор x датог канала $\hat{A}$ произвољан, па је капацитет C највећа могућа вредност излазне норме $\|\vec{y}\|$, односно могућности излазног низа.

Саму једначину $I(X;Y) = H(Y) - H(Y|X)$, њене делове, треба редефинисати. Обим „предлога“ тих редефиниција (мојих, необјављених) превелик је, претрпан сам непровереним „опцијама“, попут дезинформација. Неке идеје већ сам изнео у тексту, а ево још једне.

Грубо, узајамна информација је скаларни производ вектора, који је иначе већ уграђен у варијанси, коваријанси, или корелацији, $I \to \sum_i a_i b_i$, где су фактори разлике $x_i - \mu_x$ односно $y_i - \mu_y$. Ако је условна тзв. ентропија $H(Y|X)$ конволуција, имаћемо (отприлике) ово:

$$I(X;Y) + H(Y|X) \to \sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n} a_{n-i} b_i = \sum_{i=1}^{n} (a_i + a_{n-i}) b_i =$$

$$= \sum_{i=1}^{n} (x_i + x_{n-i} - 2\mu_x)(y_i - \mu_y) = 2\sum_{i=1}^{n} \left(\frac{x_i + x_{n-i}}{2} - \mu_x\right) \Delta y_i$$

$$\approx 2\sum_{i=1}^{n} \sigma_x \Delta y_i \approx 2\sigma'_x \sum_{i=1}^{n} \Delta y_i = 4\sigma'_x (y_1 + \dots + y_n - n\mu_y) \approx 2n\sigma'_x \sigma'_y.$$

Низови су растући, па је приви збир максимум а други минимум могућих пермутација.

59. Периодични сигнали

Посматрајмо матрицу ротације за $60°$, тј. $\frac{\pi}{6}$ радијана:

$$\hat{A} = \begin{pmatrix} \cos\frac{\pi}{3} & -\sin\frac{\pi}{3} \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

Њен квадрат је ротација за двоструки угао ($120°$), куб за троструки ($180°$) и редом даље све већег угла до шестог степена када се направи пуни круг ($360°$). Степенујући даље ротације иду као да их пре није било, из почетка, јер $\hat{A}^6 = \hat{A}$. То је пример периодичне појаве.

Слично се може десити у Марковљевом[14] ланцу или процесу[15]. Отуда нам је позната теорема да ће два комутативна (то нама значи и независна) сигнала, ако су периодични, онда бити периодични са истом периодом. Када први сигнал у n-том кораку пређе у други, а други у m-том кораку пређе у први, онда су бројеви конгруентни по модулу периода сигнала.

Нећу преписивати добро познате доказе ове теореме, јер ми се чини очигледном. Наиме, матрица канала попут наведене ротације, све периодично мења на шта год да делује. Све делове процеса које генерише она сваких шест корака понавља, али неки у то време могу бити понављани и више пута, на пример два или три. У два циклуса матрице, процес који се у сваком поновио два пута, је поновљен четири пута, а онај бржи који се у сваком циклусу понови три пута, поновиће се шест пута. Да би се таква поклапања догађала, битно је да су овде 2 и 3 делитељи броја 6.

Међутим, сигнал не мора прећи у неки други сигнал (истог ланца) да би ланац био периодичан. Следећи пример може представити разне такве могућности.

Пример 1. Низ три броја $(a, b, c) \to (b, a, c + d)$ трансформише се тако да у сваком кораку прва два замене места, а трећи се увећа за константан број d. Ако су бројеви a, b, c следећи декадни којима се води само задња цифра, а $d = 4$, генеришу ланац трансформација:

$$\begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} \to \begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix} \to \begin{pmatrix} 4 \\ 2 \\ 9 \end{pmatrix} \to \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix} \to \begin{pmatrix} 4 \\ 2 \\ 7 \end{pmatrix} \to \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix} \to \cdots \to \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$$

Након 10 корака понавља се почетни распоред. При томе горња (прва) два броја прелазе један у други, а доњи (трећи) никада не узима вредност неког од та два. ❑

У наведеном примеру, када су бројеви цели, сваки целобројни додатак генерисаће неки ланац који се понавља. Тада је доњи број увек нека од десет цифара 0, 1, ...,9 и након највише десет корака она се мора поновити. Ако су тада горње две цифре замениле места, након следећих 10 корака оне ће их опет заменити и имаћемо почетни распоред. Али ако су бројеви a, b, c цели, а

---

[14] Марковљев ланац или процес је стохастички модел који описује низ могућих догађаја у којем вероватноћа сваког догађаја зависи само од стања постигнутог у претходном догађају.

[15] Марковљев процес је верзија Марковљевог ланца са континуираним временом.

додатак ирационалан, $d \notin \mathbb{Z}$, онда нема периодичности. Ово се може посматрати још општије. Са вероватноћама је слично, само је мање извесно.

Нека је у датом Марковљевом ланцу $k_{ij}(n)$ вероватноћа да сигнал e_i пређе у неки (нови или исти) сигнал e_j након тачно n корака. Нека је $q_{ii}(n)$ вероватноћа да e_i пређе у исти сигнал након тих n корака. Збир

$$Q_i = \sum_{n=1}^{\infty} q_{ii}(n)$$

је вероватноћа да сигнал e_i у било којем кораку $n = 1,2,3, \dots$ постане опет сигнал e_i. Кажемо да је сигнал e_i повратан када је $Q_i = 1$.

Када се након n корака деси понављање $e_i \to e_i$, такво би се могло дешавати опет сваких следећих n корака и бесконачно много пута, јер оно значи већу извесност циклуса те коначне дужине. Онда то значи да ред $\sum_{n=1}^{\infty} k_{ij}(n)$ дивергира (збир је неограничен). Са друге стране, ако се то не дешава, тај ред не дивергира, онда је $Q_i = 1$, јер је низ $q_{ii}(1), q_{ii}(2), q_{ii}(3), \dots$ расподела вероватноћа. Тако можемо покушати интуитивно разумети следећи став.

Теорема 1. За сигнал e_i у хомогеном Марковљевом ланцу вреди једнакост $Q_i = 1$ ако и само ако дивергира ред $\sum_{n=1}^{\infty} k_{ij}(n)$.

Доказ: Означимо са $q_{ij}(m)$ вероватноћу да сигнал e_i након тачно m корака по први пут постане сигнал e_j $(i \neq j)$. Према формулама потпуне вероватноће имамо:

$$k_{ij}(n) = \sum_{m=1}^{n} q_{ij}(m) k_{jj}(n-m),$$

где је $k_{jj}(0) = 1$ и $n = 1,2,3, \dots$ Отуда је:

$$\sum_{n=1}^{r} k_{ij}(n) = \sum_{n=1}^{r} \sum_{m=1}^{n} q_{ij}(m)\, k_{jj}(n-m) =$$

$$= q_{ij}(1) k_{jj}(0) +$$

$$+ q_{ij}(1) k_{jj}(1) + q_{ij}(2) k_{jj}(0) +$$

$$\dots$$

$$+ q_{ij}(1) k_{jj}(r-1) + q_{ij}(2) k_{jj}(r-2) + \dots + q_{ij}(r) k_{jj}(0)$$

$$= \sum_{n=1}^{r} \left[q_{ij}(m) \sum_{n=m}^{r} k_{jj}(n-m) \right].$$

Додамо ли унутрашњем збиру овог резултата неколико ненегативних бројева, биће

$$\sum_{n=1}^{r} k_{ij}(n) \leq \sum_{m=1}^{r} q_{ij}(m) \left[\sum_{n=1}^{r} k_{jj}(n) + 1 \right].$$

Одузмемо ли истом резултату неколико сабирака биће

$$\sum_{n=1}^{r} k_{ij}(n) \geq \sum_{m=1}^{r'} q_{ij}(m) \left[\sum_{n=1}^{r-r'} k_{jj}(n) + 1 \right],$$

где је $r' < r$. Када поделимо последње две неједнакости истим збиром, ставимо да $r \to \infty$, затим $r' \to \infty$, добијамо

$$\sum_{m=1}^{r} q_{ij}(m) = \lim_{r\to\infty} \frac{\sum_{n=1}^{r} k_{ij}(n)}{1 + \sum_{n=1}^{r} k_{jj}(n)}.$$

Отуда је

$$1 - \sum_{m=1}^{\infty} q_{ii}(m) = \lim_{r\to\infty} \frac{1}{1 + \sum_{n=1}^{r} k_{ii}(n)},$$

тј. вреди једнакост $1 - Q_i = 0$ ако и само ако ред $\sum_{n=1}^{\infty} k_{ii}(n)$ дивергира. ■

Када се у неком (n-том) кораку хомогеног Марковљевог ланца појави сигнал (e_i) који је повратан ($Q_i = 1$), онда се може очекивати да ће се сваких коначно много корака појављивати исти сигнал. Такав сигнал ће се у бесконачном броју корака, са вероватноћом 1 јавити бесконачно много пута. Ово је непосредна последица хомогености ланца на основу које можемо доказати да је повратан сигнал уједно и периодичан.

Важи и обратно, ако се у неком кораку хомогеног Марковљевог ланца појави сигнал (e_i) који није повратан ($Q_i < 1$), онда можемо интуитивно предвиђати да ће се тај сигнал у остатку ланца јавити највише коначно много пута. Другим речима, тај се сигнал (e_i) са вероватноћом 1 након коначно много корака више не појављује. Управо то ћемо сада доказати.

Теорема 2. Ако за сигнал e_i у хомогеном Марковљевом ланцу вреди неједнакост $Q_i < 1$, онда појављивање тог сигнала највише коначно много пута у том ланцу има вероватноћу 1.

Доказ: Нека је ν_n број појављивања сигнала e_i у ланцу након n корака. Према претпоставци је

$$\Pr(\nu_1 < \infty) = Q_i < 1.$$

Због хомогености је

$$\Pr(\nu_2 | \nu_1) = Q_i$$

а отуда:

$$\Pr(\nu_2 < \infty) = \Pr(\nu_1 < \infty) \Pr(\nu_2 | \nu_1) = Q_i^2$$

и уопште

$$\Pr(\nu_n < \infty) = Q_i^n$$

за $n = 1,2,3, \ldots$, па је

$$\sum_{n=1}^{\infty} \Pr(\nu_n < \infty) = \sum_{n=1}^{n} Q_i^n < \infty.$$

Нека је A_n догађај да се у n-том кораку ланца појавио сигнал e_i, па је $\bigcup_{n=m}^{\infty} A_n$ догађај реализације бар једног од A_m, A_{m+1}, A_{m+2} ..., па ако то важи за свако m биће $\bigcap_{m=1}^{\infty}(\bigcup_{n=m}^{\infty} A_n)$ догађај да се реализује бесконачно много догађаја из низа $A_1, A_2, A_3, \ldots$. Отуда је:

$$\Pr\left[\bigcap_{m=1}^{\infty}\left(\bigcup_{n=m}^{\infty} A_n\right)\right] = \lim_{m\to\infty} \Pr\left(\bigcup_{n=m}^{\infty} A_n\right) \le \lim_{m\to\infty} \sum_{n=m}^{\infty} \Pr(A_n) \le \lim_{m\to\infty} \sum_{n=m}^{\infty} \Pr(\nu_n < \infty) = 0.$$

Добили смо да се вероватноћом 0 сигнал e_i бесконачно много пута јавља у Марковљевом ланцу, тј. да појавасигнала e_i коначно много пута у ланцу има вероватноћу 1. ■

Ово су познате теореме са којима сам се пре 3-4 деценије охрабривао да наставим са „теоријом информације“ (новом). Писао сам тада отприлике овако. Како је могуће да теорија информације буде формално једнака некој теорији случајних система? Само наше схватање појма „догађај“ или „стање система“ објашњава такву аналогију. Ако се десио неки догађај ми то сазнајемо преко чула, помоћу мерних апарата, саопштења итд, дакле на овај или онај начин али увек помоћу података, или информација о том догађају. Не постоји други начин да доживимо стање било каквог система сем као информацију о том систему, претпоставка је.

То онда значи да је сваки модел који правимо ради схватања стварности могућ тек онда ако је могућ одговарајући информатички модел, без обзира да ли смо овог последњег свесни. Са друге стране, ако је стварност за нас само оно што (на било који начин) неко од нас може доживети преко чула, тада је и информација стварна. Према томе је информација догађај, или му је бар еквивалентна у математичком смислу.

Брзо се затим стиже до „принципа минимализма информације“, о спонтаној тежњи природе за мањом комуникацијом, мањој информацији, еквивалента начелној природној тежњи да реализује вероватније исходе чешће. Вероватнији догађај мање је информативан. Корак од тога је принцип најмањег дејства, иначе у теоријској физици одавно примећен и кориштен. Збуњивала ме је била једино ентропија по Шенону, док се нисам одважио постулирати да већа ентропија супстанце значи мању њену (макар ефективну) информацију и ствари су легле како треба.

Дакле, информација је ткиво космоса а њена суштина је неизвесност. Већа информација иде са мањом вероватноћом и обрнуто, тако да је неизвесност такође узрок и последица убрзања. Сила мења вероватноћу, па чаша остаје на столу и у следећим тренуцима, јер је такво њено стање највероватније – све док на чашу не делује нека сила и шансе промени.

На пример, електрони у узајамном мировању привлаче се (Current), а када се покрену иду један ка другоме, крећу се у супротним смеровима и – одбијају се. Сила мења вероватноће! Стога, мењање убрзања мењаће вероватноће, ове објасниће силе, а оне тежњу ка мањој информацији. Отуда све осцилације, али и свет уопште – требало је посложити.

60. Реципроцитет

Сагласност принципу најмањег дејства је пасивност, одсуство супротстављања, препуштање као кладић низ воду. Супротност му је активност, сила, реципроцитет. Обоје обухвата „информација перцепције“, надам се, којој би требала бити блиска овдашња „узајамна информација“

$$S(\vec{a},\vec{b}) = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

За почетак за коефицијенте ових вектора узимамо само ненегативне реалне бројеве, нормиране на јединицу и, штавише, оне који представљају расподеле вероватноћа. Нагласићу који се, зашто и како, од општих налаза евентуално не односи на поједине остале врсте бројева.

Пример 1. Када се први низ раставља на збир два низа, $\vec{a} = \vec{u} + \vec{v}$, информација перцепције се раставља на збир две одговарајуће информације перцепције. Наиме:

$$S(\vec{a},\vec{b}) = S(\vec{u}+\vec{v},\vec{b}) = (u_1+v_1)b_1 + (u_2+v_2)b_2 + \cdots + (u_n+v_n)b_n =$$

$$= (u_1b_1 + u_2b_2 + \cdots + u_nb_n) + (v_1b_1 + v_2b_2 + \cdots + v_nb_n)$$

$$= S(\vec{u},\vec{b}) + S(\vec{v},\vec{b}).$$

То је особина адитивности. Хомогеност, $S(\lambda\vec{a},\vec{b}) = \lambda S(\vec{a},\vec{b})$, је такође лако доказати, па кажемо да информација перцепције има особину линеарности. ❑

Адитивност информације перцепције је разлог да је тачан и став из претходног наслова о збиру канала [58. Теорема 2.], у случају да је сматрамо узајамном информацијом. Узгред приметно је, ако $S(\vec{a},\vec{b})$ мери узајамну информацију канала, онда је она већа када јачи сигнал преноси у јачи одговарајући, а слабији у слабији. Када се догађа обрнуто, да јачи сигнал иде у слабији а слабији улазни на одговарајући јачи излазни, снага је узајамне информације канала најмања. То показује следећи пример.

Пример 2. За оба растућа (опадајућа) низа број $S(\vec{a},\vec{b})$ имаће максималну вредност у односу на све пермутације датих низова $\vec{a} = (a_1, a_2, \dots, a_n)$ и $\vec{b} = (b_1, b_2, \dots, b_n)$.

Наиме, када је $n = 2$, из очигледног $(a_2 - a_1)(b_2 - b_1) \ge 0$ следи $a_1b_1 + a_2b_2 \ge a_1b_2 + a_2b_1$. Из ненегативног резултата множења две позитивне (или обе негативне) разлике, следи тврђење. А отуда, када имамо дуже низове и први је монотон (растући или опадајући), при чему други није, замена места пару по пару другог који не следе монотоност, сваки пут ће повећавати S до највеће вредности када ће други постићи уређеност првог. ❑

На начин овог примера открива се и друга крајност, да ће број $S(\vec{a},\vec{b})$ бити минималан када су дати низови супротне монотоности, један растући а други опадајући. Ова друга крајност, најмање снаге канала, рецимо у случају једнаке уређености низова (да су оба растућа, или оба опадајућа) одговара „конволуцији“

$$S(\vec{b}|\vec{a}) = a_1b_n + a_2b_{n-1} + \cdots + a_nb_1.$$

Векторе (расподеле) намерно растављам знаком који подсећа на условне вероватноће, јер већа одређеност $\vec{a}$ смањује шансе $\vec{b}$, односно број $S(\vec{b}|\vec{a})$. Овакав збир нека је средња вредност, чини је производ просечних могућности утицаја ова два низа:

$$S(\vec{a},\vec{b}) + S(\vec{b}|\vec{a}) =$$

$$= (a_1b_1 + a_2b_2 + \dots + a_nb_n) + (a_1b_n + a_2b_{n-1} + \dots + a_nb_1)$$

$$= a_1(b_1 + b_n) + a_2(b_2 + b_{n-1}) + \dots + a_n(b_n + b_1)$$

$$\approx (a_1 + a_2 + \dots + a_n)2\mu_b = 2n\mu_a\mu_b$$

где је μ_b врста средње вредности низа $\vec{b}$, аритметичке средине одговарајућих његових најмањих и највећих бројева, а μ_a је аритметичка средина низа $\vec{a}$, док је n број чланова сваког од два низа.

Средња вредност овог збира, максимума и минимума „перцепције“, или њихове „узајамности“, расте са бројем компоненти n низова. То се противи начелном „минимализму информације“, или што је исто „максимализму вероватноће“, па се вероватно не догађа дивергенција резултата. Да би се то избегло, $2n\mu_a\mu_b \to \infty$, биће тешко постизати све већи број преносних сигнала (n), или бар већих просечних интензитета перцепција (μ_a и μ_b), или обоје. Прво ограничавање је, на пример, у броју чула индивидуе која „опажа око себе“, а друго просечне „снаге“ њених перцепција.

Видели смо да је узајамна информација канала, овако дефинисана, још једна врста информације перцепције субјекта и објекта ($\vec{a}$ и $\vec{b}$). Са већом „пропусношћу“ таквог „канала“, субјекат има бољу комуникацију са објектом, својим окружењем. Можемо рећи да је субјекат тада виталнији, упућен на више могућности (опције без могућности дејства би биле илузорне), живахнији и свестранији. Отуда радозналост и разноврсност, јер информација је са једне стране мера количине могућности, опција, а са друге еквивалент је физичког дејства. Она већа значиће већу моћ пркошења принципу минимализма теоријске физике, такође, те утолико мању мртву твар физике.

Међутим, витално стање је напето, силовито, оно тежи да се смири и да се смањи, да се препусти инерцији, лењствовању, да буде налик смрти предајући вишак информације околини, колико му то закон одржања информације дозвољава, као и околина која би и сама то исто. Алтернативна је могућност предаја идивидуалних слобода (чит. информације) вишој организацији, када се таквој може јављати емергенција, све већа виталност колектива.

Свака индивидуа има неки „појас удобности“, распон властите „узајамне информације“ који иде од сопствене минималне до максималне вредности, ограничене пре свега законима одржања. Та се величина може „напрезати“, што је видљиво рецимо код живих бића у сукобима. Вођена тиме је моја класификација „снаге игре“ (Win Lose) у три категорије игара на победу. Виталнија игра има већу узајамну информацију опонената, шта год да су они.

Гомилање информације у индивидуама може се дешавати упркос начелног минимализма, као што се слично дешава у олујама за чије честице као да не важи принцип најмањег дејства физике. То се догађа иако нема трајекторије честице која не следи из Ојлер-Лагранжових једначина, чија есенција је управо принцип најмањег дејства. Слично је и са гејзирима или вулканима за чије ерупције као да не важи привлачна сила гравитације.

Вишак виталности је пркошење, веће од онога које би чинила мање витална индивидуа, а свакако веће од мртве твари физике. У том смислу је вишак живахности вишак силе, поред вишка опција и интеракција. А како силу у односу на нешто можемо препознавати по промени брзине у односу на то, кажемо да сила (телима, честицама, таласима) даје убрзање. Зато убрзања откривају присуство сила, а то опет значи да трагом убрзања налазимо мењање (перцепције) вероватноће, те њихово повратно деловање у измени саме силе која их узрокује.

Тако можемо описати осциловање. Оно настаје сменама привлачних и одбојних сила које једна другу редефинишу. Поменуо сам да се електрони у кретању истим правцем и смером, као и у узајамном мировању, вероватније привлаче (Current), да би се покренути један ка другоме (или један од другога) кретали у супротним смеровима и тада одбијали. Истоимена наелектрисања и на тај начин (поред уобичајеног) можемо видети као одбојне силе, јер ће кретања у супротним смеровима издоминирати мировања.

Еквивалентно овоме је тумачење преноса виртуалне честице (сфере) фотона са једног електрона на други, када дође до размене спинова $\pm 1/2$ електрина са спином фотона ± 1, а све у складу са законима одржања. Слично би се могло дешавати и са разменом виртуелних гравитона (још увек експериментално непотврђених) који због превеликог спина ± 2 не интерагују непосредно са овим најмањим честицама (фермионима и бозонима), него само са њиховим спојевима. Тако се дешава да тежиште честица иде за мањом информацијом, тј. већом вероватноћом, због „гравитације“, а када се са убрзањем претера, када честица изађе из свог „појаса удобности“, окреће се смер силе, јер претходна микро-сила генерише нову.

Третирајући информацију перцепције као физичко дејство, налазимо њену промену по путу, или временом, сразмерну сили. Таква је тада и одговарајућа интерпретација узајамне информације. Кратко речено, дејство је производ импулса и пута, односно енергије и времена, а најмање такво реда је величине Планкове константе ($h = 6{,}626 \times 10^{-34}$ m²kg/s). Узимајући $\Delta S = \Delta p \Delta x$, или што је исто $\Delta S = \Delta E \Delta t$, па је аналогно и са збировима таквих, биће:

$$\frac{\Delta \mathrm{S}}{\Delta x} = \frac{\Delta \mathrm{E}}{\Delta x}\Delta t = F\Delta t, \quad \frac{\Delta \mathrm{S}}{\Delta t} = \frac{\Delta \mathrm{p}}{\Delta t}\Delta x = F\Delta x,$$

јер сила на путу чини рад који је еквивалентан промени енергије ($\Delta E = F\Delta x$), а промена силе временом мења импулс ($\Delta p = F\Delta t$).

Узвраћање информације перцепције силом на саму себе, којом се она мења, периодично или не, врста је реакције на акцију, реципроцитета који не изостају чак и када наизглед нема опонената. А тај опонент увек постоји, јер је информација перцепције бинарна мера, ње нема за објекат за којег нема субјекта опажања.

61. Гранична расподела

Дата је константна матрица канала $\widehat{K}$ и посматрамо ланац трансформација расподела $\widehat{K}^n$. Ако се дати сигнал e појавио у n-том кораку нека је вероватноћа да ће се појавити и у $n+1$-ом кораку a, а ако се није појавио у n-том кораку нека вероватноћа да ће се појавити у $n+1$-ом кораку буде b. При томе је p_n вероватноћа да се e појавио у n-том кораку. Формула потпуне вероватноће даје:

$$p_{n+1} = ap_n + b(1-p_n) = (a-b)p_n + b.$$

Ставимо $c = a - b$, па израчунавамо рекурзивно:

$$\begin{aligned} p_{n+1} &= cp_n + b = c(cp_{n-1} + b) + b = \\ &= c^2 p_{n-1} + (c+1)b \\ &= c^3 p_{n-2} + (c^2 + c + 1)b \\ &\dots \\ &= c^n p_1 + (c^{n-1} + c^{n-2} + \dots + c + 1)b \\ &= c^n p_1 + \frac{1-c^n}{1-c} b \\ &= \left(p_1 - \frac{b}{1-c}\right) c^n + \frac{b}{1-c}, \end{aligned}$$

$$\lim_{n\to\infty} p_n = \frac{b}{1-c}, \quad |c| < 1.$$

Дакле, константа $p = b/(1-c)$ којој тежи p_n не зависи од p_1. Тај број (p) је гранична или финална вероватноћа за сигнал e. Када је $0 < a < 1$, биће $b < 1 - a + b$, па је $0 < p < 1$ и гранични број је заиста вероватноћа. Супротна верованоћа је $q = 1 - p = (1-a)/(1-a+b)$.

Пример 1. Општа бинарна матрица канала има коефицијенте

$$\widehat{K} = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix}$$

За први сигнал случајног вектора (коваријантног) је $k_{11} = a$ и $k_{21} = b$. Дакле, да би први сигнал имао наведену граничну вероватноћу, прва колона не сме имати коефицијент једнак нули. Из:

$$\begin{aligned} (p,q)\begin{pmatrix} a & 1-a \\ b & 1-b \end{pmatrix} &= (ap + bq, (1-a)p + (1-b)q) = \\ &= \left(\frac{ab}{1-a+b} + \frac{b(1-a)}{1-a+b}, \frac{(1-a)b}{1-a+b} + \frac{(1-b)(1-a)}{1-a+b}\right) \\ &= \left(\frac{b}{1-a+b}, \frac{1-a}{1-a+b}\right) = (p,q) \end{aligned}$$

следи да граничне вероватноће формирају сопствени вектор дате матрице. ❑

Приметимо да је вектор $\vec{v} = (p,q)$ горњих граничних вероватноћа:

$$p = \frac{b}{1-a+b}, \quad q = \frac{1-a}{1-a+b}$$

такође сопствени вектор и следећој матрици

$$\vec{v}\begin{pmatrix} a & a \\ b & b \end{pmatrix} = \vec{v},$$

што је лако проверити. Међутим, која год од ове две је $\widehat{K}$, приметимо да $\widehat{K}^n$, за $n = 2,3,4, \dots$, не мора бити ни једна од те две. Ипак је $\vec{v}\widehat{K}^n = (\vec{v}\widehat{K})\widehat{K}^{n-1} = \vec{v}\widehat{K}^{n-1} = \vec{v}\widehat{K}^{n-2} \dots = \vec{v}\widehat{K} = \vec{v}$, па је $\vec{v}$ сопствени вектор сваке $\widehat{K}^n$.

61.1. Лакши примери

Занимљиво је да степеновањем матрице произвољних позитивних коефицијената настају матрице са не већом разликом коефицијената унутар колона (врста). То је ергодичка теорема Маркова[16] по њему зазваних ланаца. Изнећу и општи доказ на крају, а пре тога погледајмо неколико посебних једноставнијих ситуација. Разрадићу их на матрицама другог реда.

Ако су колоне матрице већ једнаких коефицијената, остаће једнаких и након квадрарања:

$$\begin{pmatrix} a & b \\ a & b \end{pmatrix}^2 = \begin{pmatrix} a^2+ab & ab+b^2 \\ a^2+ab & ab+b^2 \end{pmatrix}.$$

Ако множимо матрице једнаких коефицијената у резултату су једнаки коефицијенти:

$$\begin{pmatrix} a & b \\ a & b \end{pmatrix}\begin{pmatrix} c & d \\ c & d \end{pmatrix} = \begin{pmatrix} ac+bc & ad+bd \\ ac+bc & ad+bd \end{pmatrix}.$$

Узмимо сада општији случај, матрице другог реда позитивних неједнаких у колонама:

$$\begin{pmatrix} a & d \\ b & c \end{pmatrix}^2 = \begin{pmatrix} a^2+bd & ad+cd \\ ab+bc & bd+c^2 \end{pmatrix},$$

па ако је $0 < c - d < a - b < 1$, не губимо на општости када посматрамо разлику прве колоне:

$$(a^2+bd) - (ab+bc) = a(a-b) - (c-d) <$$

$$< a(a-b) - b(a-b) = (a-b)^2 < a - b.$$

Такође се смањује и разлика друге колоне:

$$(bd+c^2) - (ad+cd) = (c-d)c - (a-b)d <$$

$$< (a-b)c - (a-b)d = (a-b)(c-d) < c - d.$$

Дакле, разлика коефицијената унутар колона смањују се. Ако је она мања у првој колони, исто радимо са другом па првом. Коефицијенти по колонама теже ка једнаким, да се ситуација сведе на претходни случај. Наставимо ли са оваквим множењима, видећемо да у граничном случају добијамо матрице другог реда једнаких колона, када је

[16] Андреј Андрејевич Марков (1856 — 1922), руски математичар.

$$(p \quad q)\begin{pmatrix} a & b \\ a & b \end{pmatrix} = (a \quad b).$$

Резултат је вектор који не памти порекло, односно матрица је црна кутија.

Ергодичка теорија (грч. Εργον *ergon* „рад“, ὁδός hodos „пут“) грана је математике која проучава статистичка својства детерминистичких динамичких система; то је проучавање ергодичности. У „теорији информације“ која није детерминистичка, ергодичност ових канала има дубље значење. Ако постоје макар мале вероватноће да би сваки сигнал могао отићи којекуда, када би све такве биле као вероватноће канала $k_{ij} > 0$, онда нема предугог памћења чак и када има веома дуге историје.

На пример, космос који памтимо, чију прошлост форензичким или најбољим научним методама можемо „откривати“ и тако „реконструисати“, уопште не мора бити оно што је стварно било. Што старије то нетачније. Тако да би се и прича о „Великом праску“ (Big Bang) могла оспоравати.

Не морамо ићи тако далеко да приметимо конвергенције „објашњења“ догађаја које не зависе од онога што се заиста дешавало, колико од преносиоца „истине“. Ток до погрешне слике прошлости може доћи од намерне или олако прихваћене лажи, као и од заблуда. У датим околностима ту су и несхватљиве истине, а уосталом, математика нас учи да постоје и истине које су недоказиве (у оквиру дате теорије, датих знања).

На сличан начин можда се тако може разумевати и понеки „синхроницитет“, појам који је смислио психолог Карл Густав Јунг за „безузрочну повезаност“, догађаја повезаних значењима без узрока, а које људи често сматрају коинциденцијама и сматрају их случајним.

61.2. Ергодичка теорема

Теорема 1. (Ергодичка) Ако је за неко $m = 1,2,3, \ldots$ сваки елеменат матрице $\widehat{K}^m$ реда s строго позитиван, тада је за свако $j = 1,2, \ldots, s$

$$\lim_{n\to\infty} k_{ij}(n) = q_j$$

где бројеви q_j, који се називају граничне (финалне) вероватноће, не зависе од $i = 1,2, \ldots, s$.

Доказ: Показаћемо да је $\min_i k_{ij}(n)$ монотоно неопадајућа функција од n, да је $\max_i k_{ij}(n)$ монотоно нерастућа функција од n и да $\max_i k_{ij}(n) - \min_i k_{ij}(n) \to 0$. Тиме ће ова теорема бити доказана на начин како је то урадио сам Марков.

Полазећи од једначина Чепмен-Колмогорова имамо

$$k_{ij}(n+1) = \sum_{l=1}^{s} k_{il}\, k_{lj}(n) \geq \min_l k_{lj}(n) \sum_{l=1}^{s} k_{il} = \min_l k_{lj}(n).$$

Дакле, $\min_i k_{ij}(n+1) \geq \min_i k_{ij}(n)$. Слично се оказује да је $\max_i k_{ij}(n+1) \leq \max_i k_{ij}(n)$. Из $\sum_{j=1}^{s} k_{ij}(m) = \sum_{j=1}^{s} k_{lj}(m) = 1$, следи

$$\sum_{j=1}^{s} k_{ij}(m) - \sum_{j=1}^{s} k_{lj}(m) = Z_{il}^{+} + Z_{il}^{-} = 0,$$

где смо разлику лево раставили на збир позитивних (Z_{il}^{+}) и збир негативних (Z_{il}^{-}) разлика.

Према услову теореме сви су коефицијенти $k_{ij}(m)$ позитивни, па је $0 < \max_{i,l} Z_{il}^{+} = d < 1$. Даље је:

$$\max_{i} k_{ij}(m) - \max_{l} k_{lj}(m) = \max_{i,j}\left[k_{ij}(m) - k_{lj}(m)\right] \leq \max_{i,l} Z_{il}^{+} = q.$$

За $n = 1,2,3,\dots$ је:

$$\max_{i} k_{ij}(m+n) - \min_{l} k_{lj}(m+n) =$$

$$= \max_{i,l}\left[k_{ij}(m+n) - k_{lj}(m+n)\right]$$

$$= \max_{i,l} \sum_{\alpha=1}^{s} [k_{i\alpha}(m) - k_{l\alpha}(m)]k_{\alpha j}(n) \leq$$

$$\leq \max_{i,l}\left[Z_{il}^{+} \max_{\alpha} k_{\alpha j}(n) + Z_{il}^{-} \min_{\alpha} k_{\alpha j}(n)\right]$$

$$= \max_{i,l}\left\{Z_{il}^{+}\left[\max_{\alpha} k_{\alpha j}(n) - \min_{\alpha} k_{\alpha j}(n)\right]\right\}$$

$$= d \cdot \left[\max_{\alpha} k_{\alpha j}(n) - \min_{\alpha} k_{\alpha j}(n)\right].$$

Одавде следи, за $r = 1,2,3,\dots$ је $\max_{i} k_{ij}(rm) - \min_{l} k_{lj}(rn) \leq d^{r} \to 0$. ■

Ергодичка теорема показује када вероватноће прелаза у великом броју корака не зависе од почетног стања. Да финалне вероватноће чине расподелу која је сопствена за дату (константну) матрицу канала видимо из следећих једначина:

$$\begin{cases} q_1 + q_2 + \cdots + q_s = 1 \\ q_j = \sum_{i=1}^{s} q_i k_{ij}, \quad j = 1, \dots, s \end{cases}$$

Прву од једначина добијамо због

$$\sum_{j=1}^{s} q_j = \sum_{j=1}^{s} \lim_{n\to\infty} k_{ij}(n) = \lim_{n\to\infty} \sum_{j=1}^{s} k_{ij}(n) = 1$$

а другу из једначина Чепмен-Колмогорова за $m = 1$, а затим преласком на лимес када $n \to \infty$.

Пример 2. Померање са одбијајућом баријером је периодична појава коју представљају матрице:

$$\hat{K} = \begin{pmatrix} 0 & 1 & 0 \\ a & 0 & b \\ 0 & 1 & 0 \end{pmatrix}, \quad \hat{K}^2 = \hat{K}(n) = \begin{pmatrix} a & 0 & b \\ 0 & 1 & 0 \\ a & 0 & b \end{pmatrix}.$$

Лако је доказати да је

$$\hat{K}^n = \hat{K}(n) = \begin{cases} \hat{K}, & n = 1,3,5,\dots \\ \hat{K}^2, & n = 2,4,6,\dots \end{cases}$$

Нити за једно $n = 1,2,3,\dots$ нису испуњени услови претходне (ергодичке) теореме. Вектор $(0,1,0)$ се смењује са вектором $(a, 0, b)$ у ланцу Маркова генерисаним матрицом $\hat{K}$. ❑

Овде услови ергодичке теореме нису испуњени, јер сваки степен дате марице ($\hat{K}^n$) имаће нула коефицијенте. У матрици ротације [59.] ни један коефицијент степеновањем не постаје нула, али увек је бар један негативан, па опет услови ергодичке теореме нису испуњени.

Пример 3. Дата је матрица ($s = 3$) случајног померања са простом баријером је:

$$\hat{K} = \begin{pmatrix} b & a & 0 \\ b & 0 & a \\ 0 & b & a \end{pmatrix}, \quad \hat{K}^2 = \hat{K}(2) = \begin{pmatrix} b & ab & a^2 \\ b^2 & 2ab & a^2 \\ b^2 & ab & a \end{pmatrix}$$

За $ab \neq 0$ су испуњени ергодички услови. Гранична расподела $\vec{q} = (q_1, q_2, q_3)$ је сопствени вектор матрице $\hat{K}$, такав да је $\vec{q} = \vec{q}\hat{K}$, уз услов $q_1 + q_2 + q_3 = 1$. Решавајући ове једначине добијамо:

$$q_1 = \frac{b^2}{a^2 + ab + b^2}, \quad q_2 = \frac{a}{b} q_1, \quad q_3 = \frac{a^2}{b^2} q_1.$$

Специјално, ако су вероватноће прелаза $q_1 = q_2 = 1/2$, тада су $q_1 = q_2 = q_3 = 1/3$. ❑

Замислимо да пласман екипе А може бити прво, друго или треће место (податак e_1, e_2 или e_3) на неком спортском такмичењу и да се извештај о пласману екипе А преноси серијски. Нека спортски извештачи лажу (тј. $k_{ii} = 0$, $i = 1,2,3$), али не претерују ($k_{13} = k_{31} = 0$) и нису пристрасни (у свим осталим случајевима $k_{ij} = \text{const.}$). У том случају имамо горњу матрицу канала за $a = b$.

Према ергодичкој теореми ће коначни извештај о пласману, након много преноса у серији, бити једнако вероватан за за сваку од позиција ($q_1 = q_2 = q_3 = 1/3$), независно од правог успеха екипе А. Након свих тих трансформација количина информација остаје иста, али њени садржаји биће препаковани у смисао који нема везе са почетним.

***

Укратко, тзв. ергодичка теорија важна је за „теорију информације“ којом се бавим. Ергодичка теорема говори о понашању копија од копија слика након дугих узастопних копирања, када, због макар и најмањих шанси за реализацију других опција, резултат сталних пресликавања поприма карактеристике канала уместо да личи на оригиналну поруку.

На пример, усмена предања наратора историјских догађаја постајаће пре одраз карактера или жеља низа преносиоца него стварне некадашње истине. Еволуција живота тежиће адаптацији околишу све више, не марећи за почетни дизајн и задате намере (креатора, ванземаљаца, ДНК научника). Од ергодичке теореме не може побећи ни еволуција свемира. Ако је тачна поставка „теорије информације“ да постоји „објективна неизвесност“, тада је и тзв. „Велики прасак“ (Big Bang), којим је наводно пре око 13,8 милијарди година настала васиона, упитан.

61.3. Управљање каналом

Оно што је мана у прилагођавању поруке преноснику постаје предност ако желимо управљати каналом. Довољно је разумети да канал треба толерисати опције и да порука тежи да постане његова сопствена. У овом случају ствари могу бити много лакше него у општем (eigenvector).

Пример 4. Знајући каква треба бити завршна расподела, рецимо $\vec{q} = (q_1, q_2, q_3)$, формирамо константни канал генерисан матрицом

$$\hat{K} = \begin{pmatrix} q_1 & q_2 & q_3 \\ q_1 & q_2 & q_3 \\ q_1 & q_2 & q_3 \end{pmatrix}$$

па ће за сваку расподелу вероватноћа $\vec{p} = (p_1, p_2, p_3)$ резултат бити:

$$\vec{p}\hat{K} = (p_1 \quad p_2 \quad p_3)\begin{pmatrix} q_1 & q_2 & q_3 \\ q_1 & q_2 & q_3 \\ q_1 & q_2 & q_3 \end{pmatrix} =$$

$$= ((p_1 + p_2 + p_3)q_1 \quad (p_1 + p_2 + p_3)q_2 \quad (p_1 + p_2 + p_3)q_3)$$

$$= (q_1 \quad q_2 \quad q_3)$$

дакле, дата завршна расподела $\vec{q}$. ❑

Срећом, за дату циљану коначну расподелу (вектор) постоје различити процеси (матрице) којима је она сопствена. Сопствених вектора $\vec{v}$ и матрица $\hat{M}$, таквих да је $\vec{v}\hat{M} = \vec{v}$, мноштва су.

Пример 5. Када су x и y вероватноће неке расподеле ($x, y > 0$ и $x + y = 1$), а $0 < |\alpha + \beta| < 1$, онда се матрици $\hat{K}$ квадрирањем уједначавају колоне:

$$\hat{K}^2 = \begin{pmatrix} x+\alpha & y-\alpha \\ x-\beta & y+\beta \end{pmatrix}^2 = \begin{pmatrix} x + \alpha x - \beta y + \alpha\beta + \alpha^2 & y - \alpha x + \beta y - \alpha\beta - \alpha^2 \\ x + \alpha x - \beta y - \alpha\beta - \beta^2 & y - \alpha x + \beta y + xy + \beta^2 \end{pmatrix},$$

јер је разлика коефицијената у првој колони $0 < (\alpha + \beta)^2 < |\alpha + \beta| < 1$, а исто је и у другој. Обе у ретцима имају коефицијенте који су неке расподеле. Ако настављамо степеновање, $\hat{K}^n \to \hat{M}$ када $n \to \infty$, па према ергодичкој теореми добијамо једнаке колоне и граничне матрице $\hat{M}$.

Приметите да обе следеће матрице, множена и резултат множења, имају једнаке коефицијенте у свакој од по две колоне

$$\begin{pmatrix} x & y \\ x & y \end{pmatrix}\begin{pmatrix} x+\alpha & y-\alpha \\ x-\beta & y+\beta \end{pmatrix} = \begin{pmatrix} x^2 + \alpha x - \beta y + xy & y^2 + \beta y - \alpha x + yx \\ x^2 + \alpha x - \beta y + xy & y^2 + \beta y - \alpha x + yx \end{pmatrix},$$

али да ни једна од њих не мора бити гранична матрица $\hat{M}$. ❑

Међутим, алгебра, као и ергодичка теорија уче нас да не морамо бити тако радикални као у овим примерима да већ првим корацима будемо код циља. Посматрајмо проблем шире, као низ места $M_1, M_2, M_3, \dots$ тако да свако има следбеника и свако осим првог има претходника. У сваком од M_n могуће је примити поруку X_n из M_{n-1} и проследити је у M_{n+1}. Низ таквих порука $X_1, X_2, X_3, \dots$ чини један Марковљев ланац.

Преносник порука из места M_{n-1} у место M_n је канал K_n такође. Вреди општи став [53. Теорема 1] а када је канал K_n исти за свако $n = 1,2,3, \dots$ тада кажемо да је Марковљев ланац хомоген. Када је вероватноћа пријема једне поруке x иста на сваком месту, тј.

$$\Pr(X_n = x) = \text{const.}$$

за свако $n = 1,2,3, \dots$, тада кажемо да је Марковљев ланац стационаран. Да би ланац био хомоген он не мора бити стационаран, али се на њега односе последње четири теореме.

Типични примери поменутих места $M_1, M_2, M_3, \dots$, на којима можемо управљати вероватноћама појављивања одређених порука су низ РТВ релеја, низ новинских агенција које преписују текст, па и деца у игри „глувих телефона“ итд. Нетипични примери су наше свакодневне одлуке и поступци. Бирајући дејства прилагођавамо се свом текућем окружењу, али и обрнуто прилагођавамо околну себи. Сигнал се адаптира каналу, утврдили смо раније, али да се и канал адаптира сигналу, треба приметити. Бирајући шта ћемо радити, ми у толико малим нијансама мењамо велики свет около, да олако превиђамо свој шири утицај.

61.4. Белманова једначина

У свим тим узајамним утицајима, сигнала и канала, важе одговарајући закони „акције и реакције“. Што је већа количина (чит. маса) захваћених опција (дејстава), то је већа реакција (инертност, или отпор), то је већи повратни ефекат (мањег на веће). Многострукост је принципијелна особина ове теорије, а она се огледа и у различитости реакција. У том смислу је и изостајање интеракције, због чега сматрамо неће свако са сваким комуницирати.

У неким је ситуацијама технички могуће поставити канале тако да вероватноће преноса k_{ij} (i-те поруке у j-ту) зависе од неког параметра ξ, тј. $k_{ij} = k_{ij}(\xi)$, којим се опет може управљати. У том случају онај који управља параметром ξ могао би на n-том месту (M_n) повећати вероватноћу појаве унапред задате поруке упркос почетног податка. То је сложенији ниво управљања ланцем од простог избора константних матрица канала.

Општи проблем таквог управљања састоји се у следећем. Наћи вредност параметра

$$\xi = \xi(x_i, m) \quad i = 1,2, \dots \quad m = 0,1, \dots, n-1$$

као функције корака (места) m и поруке x_i на том тако да са највећом вероватноћом на n-том кораку имамо поруке

$$\pi = \{\pi_1, \pi_2, \dots\}.$$

Функција $\xi = \xi(x_i, m)$ зове се функција одлучивања, а она вредност функције $\xi^0 = \xi(x_i, m)$ за коју

$$P^0 = \max_{\xi} P(x_i \in \pi), \quad i = 1, \dots, s$$

је функција максималног одлучивања.

Означимо са $P_m(x_i)$ вероватноћу да се сигнал x_i појавио у m-том кораку. Претпоставка је да ове вероватноће зависе од функције одлучивања ξ. На основу формула потпуне вероватноће имамо

$$P_m(x_i) = \sum_j k_{ij}(\xi) P_{m+1}(x_j).$$

Одавде се добија тзв. Белманова (Bellman) једначина

$$P_m^0(x_i) = \max_{\xi} \sum_j k_{ij}(\xi) P_{m+1}(x_j), \quad i = 1,2,\dots, \quad m = 0,1,\dots,n-1.$$

Када за $m+1$ корак нађемо $\xi^0(x_j, m+1)$ и $P_{m+1}^0(x_j)$ Белманова једначина показује како се за претходни, m-ти корак одређује $\xi^0(x_i, m)$.

Ако је оператер параметра ξ јединствен тада још увек можемо имати Марковљев ланац. Међутим, када свако од места M_n канала управља параметрима следећег, онда више не морамо имати ланац са меморијом првог реда. Такав ланац није Марковљев, а његови генератори (канали) могу имати меморију.

Знамо да се свака матрица може раставити на неке факторе, $\hat{K} = \hat{A}\hat{B}$, на производе матрица од којих свака може бити врста канала, а то на више начина. Штавише, јединичне матрице се такође тако растављају, од производа инверзних матрица, преко идепотентних, па до цикличних попут ротације, рецимо за угао $60°$ чији ће сваки шести степен ($n \times 6 \times 60° = n \times 360°$) бити јединична матрица. Да такви производи могу бити растављиви у независне факторе говоре нам разни знани примери око нас, а за крај навешћу и следећи који је нетипичан.

Пример 6. Тело ротира око центра гравитације мењајући вектор брзине. Интензитет брзине редом се смењује са већом (перихел) и мањом (афел) док тело прати елипсу, да би пуна револуција тела била праћена једном пуно ротацијом ($360°$) смера брзине. О тој усклађености говоре нам теореме доказиване раније (Прилози, Централно кретање I и II), да централне константне силе опадају са квадратима удаљености и да покрећу набоје по коникама (елипсе, параболе, хиперболе), ако им се дејства шире брзином светлости.

Другим речима, исте усклађености нема у прејаким гравитационим пољима у центрима галаксија, или рецимо Меркура због близине Сунца – чији перихел (велика оса елипсе) такође кружи током Меркурових обилазака. Ајнштајнова општа теорија релативности коначно је објаснила аномалију закривљењем простора времена, а иста је изведива из принципа најмањег дејства (Минимализам Информације, Ајнштајнове опште једначине), па је онда[17] објашњива и „информатички“.

Наиме, за разлику од других сила, гравитациона продире кроз слојеве времена. То је веома слаб утицај „простора који памти“ (300 милиона пута, па и на квадрат, пута слабији), али који долази до изражаја у средини веома јаког поља. Довољно да наруши опадање са квадратом удаљености те силе и, доследно томе, кретање по коникама набоја на које делује (в. такође AGC 114905). ❑

[17] Принцип најмањег дејства, принцип најмање информације и највеће вероватноће – еквиваленти су у (мојој, још увек незваночној) „теорији информације“.

VI Део: Процеси

Када говоримо о „стохастичком процесу“ то значи да имамо неки систем и запажања у одређеним временима, а да је исход, односно посматрана вредност у сваком тренутку случајна променљива. Стохастички процес је колекција или скуп случајних променљивих индексираних променљивом t, која обично представља време.

Многима стохастичким процесима просечна својства независна су од тога где су формирани дуж временске осе. На пример, средња вредност може бити независна од t. Још једно важно својство неких стохастичких процеса је „ергодичност“, што значи да се просеци вредности неког одређеног времена попут $E[f(t)]$, могу заменити просеком током времена на било којој функцији $f(\mu_0, t)$ из стохастичког ансамбла (са фиксним μ_0). Другим речима, стохастички процес има ергодички режим ако је просек опсервабилног ансамбла једнак временском просеку. У таквом режиму, сваки скуп насумичних узорака из процеса мора представљати просечна статистичка својства целог режима.

Међу најважније спадају „максимални“, који носе највећу информацију. У случају ограничености на коначан распон исхода таква је једнолика расподела вероватноћа, у случају датог очекивања то је експоненцијална расподела, а у случају дате дисперзије то је нормална расподела. Занимљиво је да су најинформативнији канали преноса у неком смислу увек попут „црне кутије“, а да природа настоји правити отклон од таквих – због принципијелног минимализма информације. Резултат је пад обима опција, уређеност, јасноћа, памћење.

Тако можемо разумети „чудне“ резултате анализе фреквенције слова у текстовима, разних азбука, језика, од мање до више смислених садржаја. Корак од тога је схватање закона великих бројева, па чак и Дирихлеовог принципа, начином (моје) теорије информације.

62. Марковљеви процеси

Уопште говорећи, процес је промена која се догађа под утицајем времена. Марковљев процес је случајни ток зависан од времена са особином да је будућност независна од прошлости, с обзиром на садашњост. Сматра се стохастичком аналогијом диференцијалним једначинама и рекурентним односима.

Када је временска промењива Марковљевих процеса дискретна, $t \in \mathbb{N}$, називамо их дискретним временским Марковљевим ланцима. Њихове претежне алатке су основе вероватноће и линеарна алгебра, а понешто новога наћи ћете у мојој књизи „Физичка информација“, нарочито у одељцима о случајним корацима. Али, када је простор дискретан и време непрекидно, $t \in [0, \infty)$, уместо тог назива, боље им стоји назив ланци Макрова непрекидног времена. Својство Маркова, да се стања не памте, имплицира да процес има експоненцијалну расподелу (Memoryless Property). Пуасонов процес је типичан пример ланца Маркова са континуираним временом.

Нека је π скуп елементарних догађаја, а t непрекидни параметар. Стохастички процес дефинише се као функција два аргумента

$$\xi(t) = \varphi(x, t), \quad x \epsilon \pi.$$

За сваку вредност параметра t је $\varphi(x, t)$ је функција од x и случајна промењива. За сваки фиксан догађај x, функција $\varphi(x, t)$ зависи само од t и назива се реализација стохастичког процеса $\xi(t)$. То је формулација стохастичких процеса два руска математичара, А. Х. Колмогорова и А. Хинчина, око 1930. године у нализама хомогених промена које се од тада називају Марковљеви процеси.

Непосредну везу процеса и преноса података добијамо када кораке Марковљевог ланца задамо стањима у низу временских тренутака. Тада је низ случајних промењивих $X_1, X_2, \dots$, редом ознаке $X(t_1), X(t_2), \dots$, низ стања једне поруке, $X = x$, на излазима из серије канала, која се дешавају у тренуцима $t_1, t_2, \dots$. Другим речима, почетна порука $X(t_0) = X_0$ након n-тог канала је $X(t_n) = X_n$. Таква интерпретација је пожељна када радимо са великим бројевима $n \in \mathbb{N}$ и када се намеће да скуп $\{t_n\}$ посматрамо као приближно непрекидан.

Друго ограничење теорија процеса поставља код почетних услова. Када претпоставимо да је скуп елементарних догађаја π коначан (комада s), тада имамо коначну матрицу канала $\hat{K}(\Delta t)$ реда s која генерише хомоген Марковљев ланац. Даље претпостављамо да је ова матрица приближно јединична, у смислу да је

$$\lim_{\Delta t \to 0} \hat{K}(\Delta t) = \hat{I},$$

тј. $k_{ij}(0) = \delta_{ij}$ за свако $i, j = 1, \dots, s$. Када ставимо $t = 0$ за почетно време, онда је овај лимес оправдано назвати почетни услов. Показује се да је он довољан да сви коефицијенти k_{ij} као функције параметра, времена t, буду непрекидни по том параметру.

Када је време потребно сигналу за један корак Δt, дакле за пролазак кроз само један канал, тада за n корака ланца сигналу треба $t = n \cdot \Delta t$ времена. Производ n претходних стохастичких матрица опет је нека стохастичка матрица:

$$\hat{K}(t) = \hat{K}(n \cdot \Delta t) = \left[\hat{K}(\Delta t)\right]^n.$$

Посебно је:

$$\widehat{K}(t+\Delta t) = \widehat{K}(t)\cdot\widehat{K}(\Delta t) = \widehat{K}(\Delta t)\cdot\widehat{K}(t).$$

Означимо са $\widehat{D}(\Delta t) = \widehat{K}(\Delta t) - \hat{I}$. Ово је опет матрица реда s, али није стохастичка. Лако се уверити да је збир коефицијената произвољног ретка ове матрице нула. Интуитивно матрицу $\widehat{D}$ можемо схватити као матрицу преноса дезинформација. Наиме, она има само оне коефицијенте канала непромењене који i-ту поруку преносе у другачију j-ту.

Одузимањем горње једнакости добијамо:

$$\widehat{K}(t+\Delta t) - \widehat{K}(t) = \widehat{K}(t)\cdot\widehat{D}(\Delta t) = \widehat{D}(\Delta t)\cdot\widehat{K}(t).$$

Како су ове матрице коначног реда s, због горе наведеног почетног услова, разлика на левој страни једнакости тежи нули када $\Delta t \to 0$. Другим речима, сви коефицијенти су непрекидни по параметру t.

За тако велике бројеве n, када параметар t постаје приближно непрекидан, матрица $\widehat{D}$ постаје нула матрица и стога бескорисна. Зато уводимо и матрицу густина дезинформација

$$\widehat{\Lambda} = \lim_{\Delta t\to 0}\frac{\widehat{D}(\Delta t)}{\Delta t}.$$

Лако је проверити да су коефицијенти ове матрице (λ_{ij}) константе такве да је

$$\sum_{j=1}^{s}\lambda_{ij} = 0$$

за свако $i = 1,\dots,s$. Обратно, ако је збир сваког ретка матрице $\widehat{\Lambda}$ нула, тада је и збир сваког ретка матрице $\widehat{D}$ нула, те је матрица $\widehat{K}(\Delta t) = \widehat{D}(\Delta t) + \hat{I}$ стохастичка.

Полазећи од горњих једнакости и дефиниције матрице $\widehat{\Lambda}$ добијамо:

$$\frac{d\widehat{K}(t)}{dt} = \widehat{K}(t)\cdot\widehat{\Lambda} = \widehat{\Lambda}\cdot\widehat{K}(t).$$

То су матричне диференцијалне једначине Колмогорова за процес са коначно много стања. Овде су то и диференцијалне једначине дуге серије канала, коначног реда s.

Пример 1. Када је матрица неког канала (у произвољно блиској тачки $t = 0$) тачно јединична, тада је $\widehat{D}$ нула матрица, а таква је и матрица густина $\widehat{\Lambda}$. Дакле, $\lambda_{ij} = 0$ за свако $i,j = 1,\dots,s$, па се горња једнакост своди на $\frac{dk_{ij}(t)}{dt} = 0$, или $k_{ij}(t) = \text{const}$. Узимајући у обзир почетне услове, то значи да је $\widehat{K}(t) = \hat{I}$, са произвољним $t > 0$. ❑

Мало сложенији случај од овог примера је бинарна (квадратна, другог реда) матрица густина $\widehat{\Lambda}$. Да би она била добро дефинисана довољно је да $\lambda_{12} = -\lambda_{11}$ и $\lambda_{21} = -\lambda_{22}$. Отуда $\det\widehat{\Lambda} = 0$.

Покажимо да матрица $\widehat{\Lambda}$ једнозначно дефинише процес канала $\widehat{K}$. Претходна једначина је

$$\frac{d}{dt}\begin{pmatrix} k_{11} & k_{12} \\ 1-k_{11} & 1-k_{12} \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} \\ 1-k_{11} & 1-k_{12} \end{pmatrix}\begin{pmatrix} \lambda_{11} & -\lambda_{11} \\ -\lambda_{22} & \lambda_{22} \end{pmatrix}$$

што се може краће написати

$$k_{ij}'' - (\lambda_{11}+\lambda_{22})k_{ij}' + \left(\det\widehat{\Lambda}\right)k_{ij} = 0.$$

Детерминанта у последњој загради је нула, па имамо хомогене диференцијалне једначине другог реда са константним коефицијентима чије је опште решење

$$k_{ij}(t) = C_{ij}^{(1)}e^{-at} + C_{ij}^{(2)},$$

где је стављено $\lambda_{11}+\lambda_{22} = -a$.

Четири пара константи $C_{ij}^{(1)}$ и $C_{ij}^{(2)}$ одређујемо прво из почетних услова и захтева да добијемо стохастичку матрицу. Ако је $a \leq 0$, добијамо одмах $\widehat{K}(t) = \hat{I}$. Ако је $a > 0$, тј. $\lambda_{11}+\lambda_{22} < 0$, биће $k_{ij}(0) = C_{ij}^{(1)} + C_{ij}^{(2)} = \delta_{ij}$. Отуда опште решење

$$k_{ij}(t) = \delta_{ij} - C_{ij}^{(1)}(1-e^{-at}),$$

где имамо четири непознате константе C_{ij} $(i,j = 1{,}2)$ и познату константу $a = -\lambda_{11} - \lambda_{22} > 0$. Како је:

$$k_{i1} + k_{i2} = 1 - (C_{i1} + C_{i2})(1-e^{-at}) = 1$$

то је $C_{i1} + C_{i2} = 0$ за оба $i = 1{,}2$. Последње једнакости смањују број непознатих константи на две C_1 и C_2:

$$\begin{cases} k_{11}(t) = 1 - C_1(1-e^{-at}) & k_{12}(t) = C_1(1-e^{-at}) \\ k_{21}(t) = C_2(1-e^{-at}) & k_{22}(t) = 1 - C_2(1-e^{-at}) \end{cases}$$

Изводи ових коефицијената по t, у тачки $t = 0$, дају коефицијенте матрице $\widehat{\Lambda}$. Тако упоређивањем коначно добијамо:

$$C_1 = \frac{\lambda_{11}}{\lambda_{11}+\lambda_{22}}, \quad C_2 = \frac{\lambda_{22}}{\lambda_{11}+\lambda_{22}}$$

а тиме и опште, а видимо и јединствено решење за $\widehat{K}(t)$ матрице другог реда.

У специјалном случају, када је $\lambda_{11} = \lambda_{22} = -\lambda$ имамо бинарну симетричну матрицу канала

$$\widehat{K}(t) = \begin{pmatrix} \frac{1+e^{-2\lambda t}}{2} & \frac{1-e^{-2\lambda t}}{2} \\ \frac{1-e^{-2\lambda t}}{2} & \frac{1+e^{-2\lambda t}}{2} \end{pmatrix},$$

где је $\lambda > 0$. Описани и класични Марковљеви процеси разликују се само у интерпретацији.

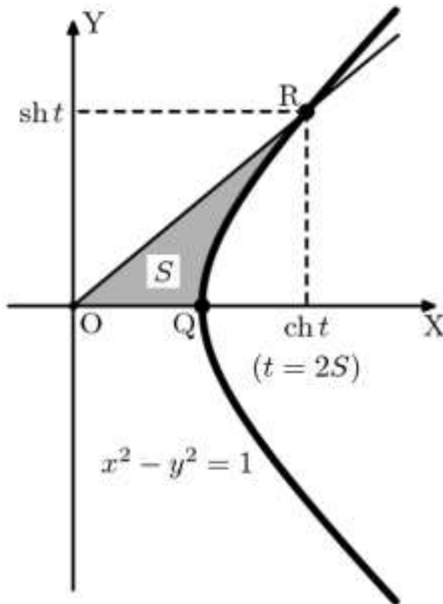
Помоћу хиперболног [04.] косинуса и синуса:

$$\operatorname{ch} x = \frac{e^x + e^{-x}}{2}, \quad \operatorname{sh} x = \frac{e^x - e^{-x}}{2},$$

ову матрицу канала пишемо

$$\hat{K}(t) = e^{-\lambda t}\begin{pmatrix} \operatorname{ch}\lambda t & \operatorname{sh}\lambda t \\ \operatorname{sh}\lambda t & \operatorname{ch}\lambda t \end{pmatrix}$$

па је можемо разумети и као хиперболну ротацију.



На слици лево је хипербола (Hyperbolic Functions)

$$x^2 - y^2 = 1$$

у Декартовом правоуглом систему координата (OXY), са потегом OR из исходишта до тачке $R(\operatorname{ch} t, \operatorname{sh} t)$, при чему је овај параметар $t = 2S$ двостука шрафирана површина, унутар криволинијског троугла RQO (између хиперболе, апсцисе и потега).

Када параметар неограничено расте ($t \to \infty$), тада потег постаје асимптота ове хиперболе, једначине праве $y = x$, а шрафирана површина достиже свој максимум ($S \to 1/2$). Тада дата матрица канала конвергира црној кутији

$$\hat{K}(t) \to \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}, \quad \lambda t \to \infty.$$

Условне вероватноће канала постају такве ($k_{ij} = 1/2$) да је информација максимална. Приметимо да је то у складу са запажањем овде [53. Теорема 1. и даље] да сметњама канал не губи информацију, већ нас само дезинформише додавањем сопственог шума.

Почетна информација се низом канала гуши, дезинтегрише или рецимом препакива, добивајући постепено нове детаље, тако да за крајњег читаоца изворна постаје све тежа за декодирати и бива коначно „изгубљена“. Када канал постане максимално оптерећен информацијом, када му условне вероватноће све постану једнако вероватне (овде су то $k_{ij} = 1/2$ за све $i, j = 1,2$) нема даље.

63. Домет сигнала

Када се у неком кораку хомогеног Марковљевог ланца појави сигнал који је повратан (поновљен, тачно једнак), онда се може очекивати да ће се сваких коначно много корака појављивати исти сигнал. Важи и обрнуто, ако се у неком кораку хомогеног Марковљевог ланца појави сигнал који није повратан, онда очекујемо да ће се тај сигнал у остатку ланца јавити највише коначно много пута [59. Периодични сигнали].

Када је бар један сигнал повратан, повратни су сви и пренос информација каналом $\hat{K} = [\![k_{ij}]\!]$ је периодичан, а бар нека од условних вероватноћа k_{ij} није строго позитиван број. Обрнуто, када бар један од сигнала није повратан, канал није периодичан и сви његови коефицијенти су строго позитивни бројеви ($k_{ij} > 0$, за све $i, j = 1, \dots, s$), а улазна информација (расподела) конвергира својственој канала, док канал постаје црна кутија [61.2. Ергодичка теорема].

Пример 1. Матрица ротације за угао $\theta = 2\pi/n$ генерише периодичан канал

$$\hat{R} = \begin{pmatrix} \cos\frac{2\pi}{n} & -\sin\frac{2\pi}{n} \\ \sin\frac{2\pi}{n} & \cos\frac{2\pi}{n} \end{pmatrix}$$

периода $n = 1,2,3, \dots$ корака. Наиме, након n корака добијамо:

$$\hat{R}^n = \begin{pmatrix} \cos\frac{2\pi}{n} & -\sin\frac{2\pi}{n} \\ \sin\frac{2\pi}{n} & \cos\frac{2\pi}{n} \end{pmatrix}^n = \begin{pmatrix} \cos 2\pi & -\sin 2\pi \\ \sin 2\pi & \cos 2\pi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

а то је идентичка трансформација. ❑

Ротација за угао $\theta = 2\pi/n$, где $n \in \mathbb{N}$, пример је „прикривања“ информације да би се појавила у изворном облику сваких n корака, ако је пренос информације периодичан. У том случају, коначног губитка информације нема, а у свакој некој (n-тој) трансформацији стохастички канал садржи по једну јединицу и све остале нуле. Таква матрица канала имаће инверзну, за разлику од матрица које имају детерминанту различиту од нуле, али се на основу примљених не могу са сигурношћу дешифровати послати сигнали [08.].

Међутим, теорија информације на којој овде радимо не подразумева постојање апсолутних нула вероватноћа, сумња у такве, па доследно очекујемо да су „периодичне“ појаве у том смислу само приближно повратне. Тако светлост која нам стиже из далеких галаксија, долази из прошлости нај даље неких 13 милијарди година. Из мало ближе прошлости видимо је са „црвеним помаком“ јер потиче са места која беже од нас то брже што су даља, односно што су старија. Још ближа светлост око нас одбија се, прелама или стапа са околним објектима, тако да њене осцилације опет не трају неограничено.

Експериментални доказ вечности физичке појаве немогућ је и, према томе, претпоставка да нема апсолутних нула условних вероватноћа евентуалних матрица канала прихватљива је. Она је у овој теорији подржана и претпоставком да је информација ткиво простора, времена и материје, а да је њена суштина неизвесност.

Тако гледајући, честица-талас (светлост) која допутује до нас кроз огромна пространства космоса нека је информација. Само њено путовање такође је информација, па се отвара питање и о успут насталој информацији трајекторије, о додатној и „наталоженој“ негде другде, ако већ честица не постаје све већа и већа на свом путу него напротив губи енеригију[18]. Али то има даље занимљиве импликације.

Доплеров ефекат удаљавања далеких галаксија може се посматрати и кроз ергодичку теорему, да преносом кроз низ канала информација постаје све више ,,својствена'' (eigen, тј. карактеристична) каналу, на уштрб оригиналне. Ми зато никаквим експериментима нити новим научним знањима можда нећемо моћи засигурно утврдити да је ,,Велики прасак'' (Big Bang) заиста постојао, или шта је тачно било пре тих 13,8 милијарди година. Пад фреквенције светлости ,,извора који се удаљава'' можда је просто још једна од последица ергодичке теореме [61.2.].

Путујући кроз низ канала, оригинална информација обогаћује се шумом. То постаје сметња њеном читању на начин као што пласт сена може сакрити иглу, или као што би додавањем до свих седам цртица дигиталних бројева на овом сату сви постали број 8 и тако „обогатили“ приказ да би он постао неразумљив.



Са мањком информације нисмо информисани, а са вишком смо дезинформисани – неко је давно приметио.

Сама „ергодичка теорема“ ипак се може третирати и као последица ,,принципа најмањег дејства'' информације, односно чешће „реализације вероватнијих исхода“ а мање информативних таквих. Наиме, она кроз канале тежи мањим променама, вишак информације настао преносом уједно је вишак дејства (дејство, производ настале енергије и при томе потрошеног времена, еквивалент је информацији), које се затим може дати околини путем познатих физичких интеракција. То рецимо значи да нама познато загревање преносника информација (радио, ТВ, телефон, компјутер) неће бити последица само трошења енергије потребног за покретање уређаја, већ је ту и загревање од протока информација. Где то дејство одлази, ако фотон путује вакуумом – ново је питање.

Додатно, обзиром да се фермиони могу распадати у бозоне и обрнуто, честице-таласи супстанце обострано у одговарајуће простора, ако су први преласци у просеку мало вероватнији од других, то је разлог плус да буде све више простора у све ређој супстанци васионе. На тај начин ширење свемира може, уз све претходно речено, бити и објективна појава.

Такође, ако „простор памти“ на описане начине, онда због закона одржања и ставова да систем са све већом ентропијом има све мању информацију (Ентропија), да се други закон термодинамике и уопште спонтаног раста ентропије, односи само на супстанцу, а не и на простор. Сва информација је сачувана у збиру и након „топљења“ супстанце у простор, она остаје константна у својој укупној количини садашњости и сећања, а за ентропију не важи закон одржања.

Моја теорија информације озбиљно узима и гравитационо деловање прошлости на садашњост, са смањеним дометом око 300 милиона пута па на квадрат (због временске координате ict), у истом систему физичких мера (метар, килограм, секунда). Тако разумемо повлачење простора за масом у јаком гравитационом пољу, рецимо померања напред перихела Меркура у ротацији око Сунца,

[18] Светлост губи фреквенцију, сматра се због Доплеровог ефекта извора који се удаљавају.

па и присуства тамне материје у галаксијама које нису изузетно ретке (AGC 114905). Тиме поново стижемо на питање домета утицаја, али сада и прошлости на садашњост, па и на „домет канала“.

63.1. Грешке копирања

Процена је да се грешка у копирању једног бита деси отприлике једном на сваких 74 TB података. Терабајт (TB) је јединица дигиталних података која износи око трилион бајтова. Тачније, 1 TB је 2^{40}, или 1 099 511 627 776 бајтова, а један бајт је 8 бита. То је грешка која се дешава и мање од двапут месечно у просечном раду кућног компјутера.

Зато сервери користе ECC (Error Correction Code) меморије. Систем са таквом меморијом типично препознаје једнобитну грешку, тако што уместо 8-битног бајта користи 9-битни, где девету цифру, 0 или 1, аутоматски додаје тако да је број јединица улазног бајта рецимо паран. Ако се тај паритет наруши, постоји грешка копирања. Затим се користе даљи начини откривања детаља и исправке грешке, или се тражи ново учитавање сумњивог податка. Шансе грешке се тим поступно смањују, увек по цену трошења меморије и времена, али не исчезавају у потпуности.

Пренос информације кроз дуге серије канала може се посматрати и тако, као узастопно копирање копија. Статистика ће се на крају задовољити вероватноћама појаве сигнала, нарочито ако их има на сваком делу дуге серије. Таква вероватноћа (p) је функција неког параметра (t), који опет може довољно тачно представљати и дужину проводника. Низ вероватноћа улазне расподеле, $\vec{p} = \vec{p}(t)$, усредњава се експоненцијалним брзином у интервалу $(0, t)$ када t расте од 0 до ∞, постају такве које канал више не мења (својствене), сагласно начелном минимализму промене информације, а канал постаје црна кутија (за полазну информацију).

Пример 2. Опште решење једначина Колмогорова за бинарну симетричну матрицу канала је

$$\hat{K}(t) = \begin{pmatrix} \frac{1+e^{-2\lambda t}}{2} & \frac{1-e^{-2\lambda t}}{2} \\ \frac{1-e^{-2\lambda t}}{2} & \frac{1+e^{-2\lambda t}}{2} \end{pmatrix},$$

где је $\lambda > 0$, као што смо видели на крају претходног наслова [62.]. При томе

$$\hat{K}(t) \to \hat{K}_\infty = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \lambda t \to \infty.$$

што значи да оригинална расподела $\vec{p}_0 = (x_0, y_0)$, постепено прелази у расподелу $\vec{p}_\infty = (\frac{1}{2}, \frac{1}{2})$, при томе се прва (већа) вероватноћа смањује експоненцијалном брзином, а друга (мања) истом расте. У том граничном случају, сваки вектор постаје $\vec{p}_\infty$, јер је:

$$\vec{p}_0 \hat{K}_\infty = (x_0 \quad y_0)\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} \frac{x_0+y_0}{2} & \frac{x_0+y_0}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \vec{p}_\infty$$

без обзира какав је улазни вектор. ❑

У општем случају, претпоставка је да i-та порука послата мрежом, у тренутку $t = 0$, припада неком случајном скупу са расподелом $\vec{p} = (p_1, \dots, p_s)$, где је p_i вероватноћа да се на улазу појави баш i-та порука. Након времена $t > 0$, вероватноћа да ћемо примити j-ту поруку је

$$q_j(t) = \sum_{i=1}^{s} p_i k_{ij}(t), \quad j = 1, \dots, s.$$

Ергодичка теорема за Марковљеве ланце даје финалне вероватноће

$$q_j^* = \lim_{t\to\infty} q_j(t)$$

за $j = 1, \dots, s$ и финалну расподелу $\vec{q}^* = (q_1^*, \dots, q_s^*)$ која може бити битно различита од почетне расподеле $\vec{p}$. Уместо да поново решавамо диференцијалну једначину Колмогорова, сада матрице реда s, погледајмо мало другачији пример, који показује да је иста зависност од сметњих излазне расподеле као и врсте канала.

Пример 3. За канал $\hat{K}(t)$ када је $\vec{q}(t) = \vec{p}\hat{K}(t)$, у општем случају за j-ту колону, опет помоћу диференцијалних једначина Колмогорова налазимо:

$$\frac{dq_j(t)}{dt} = \sum_{i=1}^{s} p_i \frac{dk_{ij}(t)}{dt} = \sum_{i,l=1}^{s} p_i\, k_{il}\lambda_{lj} = \sum_{l=1}^{s} q_l(t)\lambda_{lj},$$

$$\frac{dq_j(t)}{dt} = \vec{q}(t)\hat{\Lambda}$$

уз почетни услов $\vec{q}(0) = \vec{p}$. Дакле исту зависност од параметра сметњи $\hat{\Lambda}$ има излазна расподела као и врста канала $\hat{K}(t)$. Без обзира како дефинисали домет, сасвим је сигурно да се он значајно повећава када свако $\lambda_{ij} \to 0$ $(i \neq j)$. ❑

Због оволике важности канала у преносу поркука, корисно је да се сваки део водича састоји од канала преноса информације што је могуће приближније форме јединичне матрице, а да кôд који се преноси буде коначне азбуке. Овај други услов потребан је процесима Маркова, непосредним интерпретацијама ланаца Маркова, а занимљиво је да се може једнако примењивати и на ручно копирање дигиталних података са једне меморије на другу.

63.2. Препорука

Услов да сметња у проводнику не зависи од сигнала којег проводник носи може се записати

$$\lambda_{ij} = \begin{cases} -(s-1), & i = j \\ \lambda, & i \neq j \end{cases}$$

По дефиницији је

$$\lambda = \lim_{\Delta t\to 0} \frac{k_{ij}(\Delta t)}{\Delta t}$$

те је $\lambda \geq 0$. Константа λ представља густину вероватноће да ће произвољан i-ти сигнал, док прође кроз врло мали део проводника, изаћи као другачији j-ти сигнал. У савременој техници преноса је $\lambda > 0$ врло мали број.

Сабирци са десне стране једначина Колмогорова сада су сви множени истим бројем λ сем једног. Лако је проверити да се те једначине своде на:

$$\frac{dk_{ij}(t)}{dt} + s\lambda q_j(t) = \lambda$$

уз почетни услов $\vec{q}(0) = \vec{p}$. То је линеарна диференцијална једначина првог реда са константним коефицијентима. Њено опште решење је

$$q_j(t) = \frac{1}{s} + C_j e^{-s\lambda t}, \quad j = 1, \dots, s$$

где су $C_j = p_j - 1/s$ константе одређене почетним условима. Када $t \to \infty$, тада $q_j(t) \to q_j^* = 1/s$. Према томе, формалне вероватноће имају расподелу са максималном информацијом. У реалним условима ове су промене далеко повољније.

Због неједнакости $e^x \geq 1 + x$, биће $-e^{-s\lambda t} \leq -(1 - s\lambda t)$, тј.

$$|p_j - q_j(t)| \leq \lambda t |1 - sp_j|$$

за свако $j = 1, \dots, s$. Како је $0 \leq p_j \leq 1$, добијамо даље

$$0 \leq |p_j - q_j(t)| \leq |s - 1|\lambda t.$$

Коефицијент (λ_{ii}) уз t представља густину вероватноће грешке при преносу сигнала на основном делу проводника. Претпостављамо да је тај број врло мали. Брзине преноса сигнала могу бити толико велике (блиске брзини светлости) да је укупно време t врло мало. Време обиласка сигнала око земљине кугле, на пример, је мање од једне секунде.

Према овим релацијама имамо следећу препоруку за повећање домета, или што је исто, за бољи пренос података:

a) Бирати преносник веће брзине;
b) Смањити техничке сметње λ;
c) Смањити дужину кода s.

Последњи услов (c) јесте, али не толико битан колико на први поглед изгледа. У случају велике дужине кода, s, само мали број од свих сигнала се преноси са вероватноћом већом од нуле. Да није тако било би практично немогуће пренети поруку са непрекидном расподелом.

63.3. Неограничена азбука

Када број слова кодне азбуке $s \to \infty$, проблематици можемо прићи са другачије стране. Означимо са $F(x,t)$ вероватноћу да је сигнал π прешао пут x за време t. Претпоставимо да је са места $x = 0$ у тренутку $t = 0$ послат сигнал π. Тада су почетни услови

$$F(x,0) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases}$$

Ма како била велика брзина V медија који преноси поруку, свакако је $V < +\infty$. Отуда други, доњи од услова.

Претпоставимо даље да је p вероватноћа да ће сигнал π прећи пут h за време τ. Онда је $1 - p$ вероватноћа да се π неће појавити на излазу, па је према формулама потпуне вероватноће:

$$\begin{cases} F(x,t+\tau) = pF(x-h,t) + (1-p)F(x,t) \\ F(x+h,t) = pF(x,t-\tau) + (1-p)F(x,t) \end{cases}$$

Одузимањем добијамо:

$$\begin{cases} F(x,t+\tau) - F(x,t) = p[F(x-h,t) - F(x,t)] \\ F(x+h,t) - F(x,t) = p[F(x,t-\tau) - F(x,t)] \end{cases}$$

На основу Тејлорове формуле је:

$$F(x-h,t) - F(x,t) = -h\frac{\partial F(x,t)}{\partial x} + \frac{h^2}{2}\frac{\partial^2 F(x,t)}{\partial x^2} + o(h^2),$$

$$F(x,t-\tau) - F(x,t) = -\tau\frac{\partial F(x,t)}{\partial t} + o(\tau),$$

$$F(x,t+\tau) - F(x,t) = +\tau\frac{\partial F(x,t)}{\partial t} + o(\tau),$$

$$F(x+h,t) - F(x,t) = h\frac{\partial F(x,t)}{\partial x} + \frac{h^2}{2}\frac{\partial^2 F(x,t)}{\partial x^2} + o(h^2).$$

Сменом у претходним једнакостима добијамо:

$$\begin{cases} \tau\frac{\partial F}{\partial t} + o(\tau) = p\left[-h\frac{\partial F}{\partial x} + \frac{h^2}{2}\frac{\partial^2 F}{\partial x^2} + o(h^2)\right], \\ h\frac{\partial F}{\partial x} + \frac{h^2}{2}\frac{\partial^2 F}{\partial x^2} + o(h^2) = p\left[-\tau\frac{\partial F}{\partial t} + o(\tau)\right]. \end{cases}$$

За велике брзине V је τ толико мање од h да остаци $o(\tau)$ и $o(h^2)$ када $t \to 0$, постају једнако занемарљиви. Додатно са $h = V\tau$, где је V брзина сигнала, налазимо:

$$\begin{cases} \frac{\partial F(x,t)}{\partial t} = -pV\frac{\partial F(x,t)}{\partial x} + \frac{pVh}{2}\frac{\partial^2 F(x,t)}{\partial x^2}, \\ \frac{\partial F(x,t)}{\partial x} = -\frac{p}{V}\frac{\partial F(x,t)}{\partial t} - \frac{Vh}{2}\frac{\partial^2 F(x,t)}{\partial x^2}. \end{cases}$$

Добили смо парцијалне диференцијалне једначине попут Планк-Фокнерових за хаотично кретање честица гаса, или једначинама Колмогорова за случајне процесе непрекидне расподеле.

У првој једначини је вероватноћа да ће сигнал (π) прећи пут h, а у другој да ће након времена τ то остати исти сигнал π. Због уграђеног услова, брзине сигнала $V = h/\tau$, можемо претпоставити да су те вероватноће једнаке, чак и ако је $p = p(x)$ и $V = V(x)$. За разлику од Планкове једначине где је приближно $p = 1/2$, овде има смисла претпоставка да је приближно $p = 1$.

Када је $p = 1$ сабирањем добијамо

$$\frac{\partial F(x,t)}{\partial t} + V\frac{\partial F(x,t)}{\partial x} = 0.$$

Решење ове диференцијалне једначине је свака диференцијабилна функција $F(x,t) = f(Vt - x)$, где је $f(0) = 1$, $f(-x) = 0$ $(x > 0)$, због почетних услова. Према томе, максимална вероватноћа $F(x,t) = 1$ је достигнута у свакој тачки у којој је $x = Vt$. Домет сигнала је неограничен.

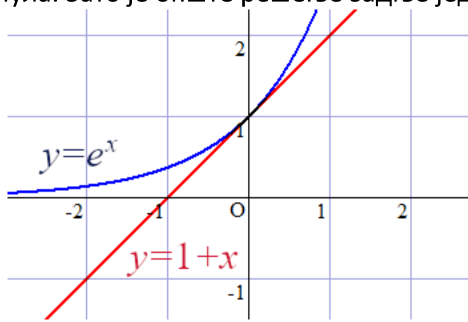
Када је $p = 1 - \varepsilon$, где је $\varepsilon > 0$ веома мали број, биће приближно $\frac{1}{p} = 1 + \varepsilon$ па одузимањем

$$\frac{\partial^2 F(x,t)}{\partial x^2} + \frac{\varepsilon}{h}\frac{\partial F(x,t)}{\partial x} = 0,$$

а отуда

$$\frac{\partial F(x,t)}{\partial x} + \frac{\varepsilon}{h} F(x,t) = \text{const.}$$

Јасно је зашто десна страна последње једнакости не зависи од x. Међутим, према претпоставци $(x = Vt)$ она није нити функција од t. Почетни услов $F(x,0) = 0$ даље тражи да ова константа буде нула. Зато је опште решење задње једначине $F(x,t) = C \cdot \exp(-\varepsilon x/h)$, што због почетног услова



$F(0{,}0) = 1$ даје константу $C = 1$. Тако, вероватноћа да ће сигнал π достићи тачку x износи

$$F(x,t) = e^{-\frac{\varepsilon x}{h}}.$$

Затим, помоћу неједнакости $e^x \ge 1 + x$ добијамо

$$\varepsilon < |1 - F| \le \frac{\varepsilon x}{h}.$$

Дакле, домет сигнала се смањује када се „сметња“ ε повећава док смањивање x (или τ) не утиче толико на домет јер се тада смањује и ε $(p = 1 - \varepsilon)$.

63.4. Аналогије

Горња функција $F = F(x,t)$ даје вероватноће да ће сигнал π прелазити путеве x за времена t. Вероватноћа p је њен посебан случај, да ће сигнал π прећи пут h за време τ, дакле $p = F(h,\tau)$, а $\varepsilon = 1 - p$ је вероватноћа да се π неће појавити на излазу. Општије, $1 - F(x,t)$ је вероватноћа да сигнал није прешао пут x за време t.

Међутим, ако је $F = \exp(-\varepsilon x/h)$ вероватноћа, онда је $\varepsilon x/h = -\ln F$ нека информација. Обзиром да су ε и h константе, то ова информација зависи само од x, односно од вероватноће $F = F(x,t)$ да сигнал π пређе пут x за време t. Ако $\varepsilon x/h$ видимо као Хартлијеву информацију (логаритам једнако вероватних исхода), онда нам је F мноштво неких (апстрактних) уједначених исхода. Затим се намеће питање поопштења на комплексне бројеве.

Отуда аналогија функције $F(x,t)$ са таласном функцијом $\Psi(x,t)$, иначе решењем Шредингерове једначине за честицу-талас квантне механике, има основе. Посебно, за слободну честицу биће

$$\psi(x,t) = Ae^{i(kx - Et)/\hbar}$$

где је A константа нормирања вероватноће, за имагинарну јединицу важи једнакост $i^2 = -1$, таласни број $k = \sqrt{2mE}/\hbar$, енергија честице $E = mc^2$, а ħ је Планкова редукована константа.

Решење Шредингерове једначине за основно стање погледајте у мојој књизи „Квантна Механика“ (Пример 1.4.5.). Добија се

$$\psi(x) = \sqrt{k}\frac{e^{k(x-x_0)} \pm e^{-k(x-x_0)}}{2}$$

у тачки x_0. Слична су и решења „потенцијалног прага“ (1.4.6 Потенцијалне препреке) када честица пролази интервалима апсцисе, први $x < 0$, други $x \in [0, a]$ и трећи $x > a$, различитих константних потенцијала као препрека

$$\psi(x) = Ae^{ikx} + Be^{-ikx},$$

где константе A и B, као и одговарајући таласни број k, зависе само од дате три позиције честице на апсциси.

Знамо да експоненцијалној расподели вероватноћа теже дискретни процеси попут Бернулијевих, рецимо бацања новчића – где исход следећег не зависи од претходног. У такве спадају и процеси Маркова, стохастички модели који описују догађаје са вероватноћама сваког од њих независима од претходних стања. Исти модел независности довољно важи и за таласе-честице које налазимо у решењима Шредингерове једначине. Куда год да такве честице иду, форма и вероватноће остају им једнаке. Све су то ситуације експоненцијалних расподела аналогних по „непамћењу“.

64. Без памћења

Бацање новчића типичан је пример расподеле вероватноће без меморије. У сваком том бацању су шансе да падне „писмо“ (вероватноће $0 \leq p \leq 1$) или „глава“ (вероватноће $q = 1 - p$) константне, без обзира шта је пре било (у случају фер новчића $p = q = 1/2$).

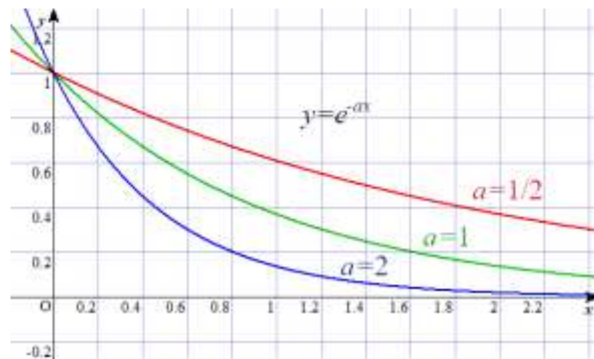
Сличан пример су независни кварови машина. Када израчунавамо вероватноћу квара (независног, новог) уређаја често није важно колико се хардвер кварио раније. Вероватноћа да ће отказати пар минута од сада не зависи од чињенице да се није кварио два месеца. Такође ће наилазак купаца у продавницу бити „безмеморијски“, када сазнање о уласку последњег од њих не помаже, а купци су неповезани, са индивидуалним понашањем.

За дистрибуцију вероватноће у статистици се каже да има својство без меморије када вероватноћа будућег догађаја не зависи од појаве прошлих. Дистрибуција вероватноћа без памћења углавном је експоненцијална, затим геометријска са природним бројевима, међу које спада и Вејбулова.

На слици десно су три графа, експоненцијалне функције

$$y = e^{-ax}$$

са параметрима $a = 1/2$, $a = 1$ и $a = 2$, редом горњи црвен, средњи зелен и доњи плав. Како је извод ових функција, $y' = -ay$, то се нагиб (однос промене ординате променом апсцисе) мења сразмерно ординати, тј. вероватноћа се мења сразмерно вероватноћи.



Још једну занимљиву особину експоненцијалних расподела видели смо у наслову „Ограничења“ [41. Теорема 1. *ii*.]: Нека је дата густина вероватноће $f(x)$ са непрекидном реалном варијаблом x. Ако је $f(x) = 0$ за $x < 0$ и ако постоји коначно очекивање μ тада постоји и информација за коју важи $S \leq \log(e\mu)$, при чему једнакост вреди ако и само ако f има експоненцијалну расподелу.

Формално кажемо да случајна промењљива X прати расподелу вероватноће са својством да је без меморије, ако је за било које m и n из {0, 1, 2, …} важи

$$\Pr(X > m + n | X \geq m) = \Pr(X > n).$$

Другим речима, ако смо имали $m = 10$ покушаја без успеха, вероватноћа да ћемо морати чекати до 30. покушаја или већег да бисмо доживели успех је иста као вероватноћа да почнемо од нуле и имамо 20 неуспелих покушаја или више до повољног исхода.

Сигнал π који наилази на датом месту (апсцисе $x \geq 0$) у датом тренутку ($t \geq 0$), са вероватноћом $F(x, t)$, било да је слово неког дугачког текста (књиге), као и реч изговорена током подужег радио програма, или аутомобил одређене боје који ће проћи раскрсницом, углавном вреди посматрати као расподеле без памћења. Тако је и са таласима-честицама квантне механике, вероватноћама њиховог налажења $\psi(x, t)$ на датом месту (x) у датом тренутку (t), које су толико елементарне да не могу носити и своје прошлости. Зато је за очекивати да оне, сигнали и честице [63.4. Аналогије], имају сличне расподеле.

Чудна је употреба комплексних бројева за такве вероватноће у квантној механици, са становишта данашње физике. Међутим, са позиција моје теорије информације они се добро уклапају у причу (Bypass). Не бих се понављао, сем да подсетим да су комплексна подручја такође неки исходи, та реална тек у комбинацијама (збировима) са себи сличнима.

64.1. Независност

Навешћу прецизније услове у којима се јављају експоненцијалне расподеле, јер се оне врло често срећу у применама и вреди те ситуације разликовати других њима наизглед сличних. Најчешће су ти услови у теорији вероватноће познати као Пуасонов процес.

Замислимо да у једној тачки проводника података, у току неког периода времена (t) региструјемо одређену групу сигнала ($\pi_1, \pi_2, \dots, \pi_s$). Оно што израчунавамо је вероватноћа појаве ових сигнала у датом периоду.

Теорема 1. (Експоненцијални процес) Када појава бар једног сигнала из скупа $\{\pi_1, \pi_2, \dots, \pi_s\}$ у временском периоду дужине $t \ge 0$ има вероватноћу $F(t)$, где је $F(t)$ непрекидна и растућа функција времена, ако је та вероватноћа независна од почетка периода t и независна од претходних сигнала, тада је $F(t) = 1 - e^{-\lambda t}$, где је параметар $\lambda \ge 0$.

Доказ: Нека је p вероватноћа да се сигнал неће појавити у периоду јединичне дужине ($t = 1$), а сам период поделимо у n једнаких (дисјунктних) интервала. Нека је $p_0(1/n)$ вероватноћа да се сигнал неће појавити у таквој једној подели ($t = 1/n$). Због услова независности је:

$$p = [p(1/n)]^n, \quad \text{или} \quad p(1/n) = p^{1/n}.$$

Отуда је вероватноћа да се сигнал неће појавити за време $t = k/n$, за $k = 1,2, \dots$, једнака:

$$p_0(k/n) = p^{k/n} \quad \text{и} \quad F(k/n) = 1 - p^{k/n}.$$

Како је $F(t)$ растућа функција времена, за $(k-1)/n \le t \le k/n$ биће:

$$F\left(\frac{k-1}{n}\right) \le F(t) \le F\left(\frac{k}{n}\right).$$

Функција $F(t)$ је и непрекидна, па можемо прећи на лимес:

$$\lim_{n\to\infty} F(t) = 1 - p^t.$$

За сваки број $p \in (0,1)$ постоји број $\lambda \in (0, \infty)$ такав да је $p = e^{-\lambda}$, чиме је теорема доказана. ∎

Последњом теоремом доказано је да независан случајни процес (без памћења) мора бити баш експоненцијалне расподеле, али из доказа се може се видети да „скоро увак" важи и обрнуто, јер важи $e^{-\lambda n} = \left(e^{-\lambda}\right)^n = p^n = p_1 \cdot p_2 \cdot \ldots \cdot p_n$ чак и када су p_i-ови различити, али ако су те једнакости (статистички) тачне за свако n та је различитост мање вероватна и важи и обрат теореме.

Вероватноћа $F(t) = 1 - e^{-\lambda t}$ има густину, по дефиницији једнаку:

$$\rho(t) = \frac{dF}{dt} = \lambda e^{-\lambda t},$$

а то нам је позната густина експоненцијалне расподеле непрекидног параметра t. Позната нам је и средња вредност (очекивање) појаве (бар једног) датог сигнала $\mu = 1/\lambda$. Пре тога смо видели и доказ још значајнијег става, да унапред дата средња вредност (очекивање) расподеле са захтевом да информација буде минимална, дају баш експоненцијалну расподелу [Информатичка теорија I, 41. Теорема 1. *ii*, 108. страна]. Ово двоје, да је експоненцијална расподела без памћења и најмање информације, у складу су са начелом штедње комуникације (интеракције).

Поновимо, ако систем отказује око $n = 1,2, \dots$ пута током T часова рада, претпоставићемо да он отказује по једном просечно сваких $\tau = T/n$ часова и да је интензитет отказа $\lambda = 1/\tau$ отказа/сат. Тада је:

$$\Pr(0,t) = \int_0^t \rho(t)\,dt = \int_0^t \lambda e^{-\lambda t}\,dt = -\left(e^{-\lambda t}\right)_0^t = 1 - e^{-\lambda t} = F(t)$$

вероватноћа да ће током времена t систем отказати. Када $t \to \infty$, тада $F(t) \to 1$.

Пример 1. Ако систем откаже 4 пута током 1000 часова рада, претпостављамо да он отказује по једном просечно сваких 1000/4 = 250 часова и да је интензитет отказа $\lambda = \frac{1}{250} = 0{,}004$ отказа/сат. Тако је:

$$F(t) = 1 - e^{-\lambda t} = 1 - e^{-0{,}004\times 500} = 1 - e^{-2} = 0{,}86466.$$

вероватноћа да ће систем отказати бар једном током $t = 500$ часова рада. ❑

Замислимо да имамо једну телефонску линију и један телефонски апарат. Поруке стижу независно и позиви имају независна трајања, тако да и почетак и трајање разговора имају експоненцијалну расподелу, са параметрима α и β редом. Интересантна су нам бар два стања таквог система:

- π_1 – телефон слободан,
- π_2 – телефон заузет.

Видели смо од раније да је $k_{12}(\Delta t) = 1 - e^{-\alpha\cdot\Delta t}$ вероватноћа преласка из стања π_1 у π_2 за неки период $\Delta t \to 0$, а да је обратно $k_{21}(\Delta t) = 1 - e^{-\beta\cdot\Delta t}$. Даље треба израчунати вероватноћу $k_{ij}(t)$ у произвољном тренутку t. Претпостављамо да се нови позив губи ако је телефон заузет, а да су π_1 и π_2 једина стања. Коефицијенте матрице $\widehat{\Lambda}$ тражимо по дефиницији:

$$\lambda_{12} = \lim_{\Delta t\to 0}\frac{k_{12}(\Delta t)}{\Delta t} = \alpha, \quad \lambda_{21} = \lim_{\Delta t\to 0}\frac{k_{21}(\Delta t)}{\Delta t} = \beta.$$

Систем диференцијалних једначина Колмогорова даће нам опште решење бинарне матрице $\widehat{K}$ које смо већ израчунали [63.1. Грешке копирања].

За бољи пренос података препоручујемо [63.2]: преносник веће брзине, мањих техничких сметњи (λ), краће кодне азбуке (s слова). Ово троје је препорука, јер управо је то интерпретација (слабо за сада примећена) решења иначе познатих једначина Колмогорова [63.2. Препорука]. Међутим је додато да овај трећи није превише битан, а како и зашто није „толико битан“, видећемо.

Укратко, трећа препорука је мање значајна, јер дуже азбуке нека слова занемарљиво ретко користе. Зато, рецимо, анализирајући сигнале из свемира, као и говор делфина, према облику графа фреквенције сигнала можемо препознати интелигентну комуникацију, чак и не знајући о чему дотични „причају“.

64.2. Садржајност

Да се уверимо да дуже азбуке нека слова занемарљиво ретко користе, када текст садржи поруку са смислом, можемо проверавати „одокативно“, али и са једноставним програмчићима анализа фреквенција слова азбука по текстовима. То су задачићи које можемо сами испробавати, а ево примера једног од резултата.

Пример 2. Из моје књиге „Приче о информацији“, узео сам поглавље „1.1. Слобода“ и пребројао велика и мала слова српске азбуке (компјутером). Од укупно 2803 знакова тог текста, 2268 била су нека од 30 великих, или исто толико малих слова српске азбуке са следећом фреквенцијом:

1. А + а = 0 + 272 = 272 → 272/2268 = 11%
2. Б + б = 0 + 50 = 50 → 2%
3. В + в = 0 + 74 = 74 → 3%
4. Г + г = 0 + 37 = 37 → 2%
5. Д + д = 3 + 100 = 103 → 103/2268 = 5%
6. Ђ + ђ = 0 + 3 = 3 → 0%
7. Е + е = 0 + 192 = 192 → 8%
8. Ж + ж = 0 + 14 = 14 → 1%
9. З + з = 4 + 51 = 55 → 2%
10. И + и = 1 + 198 = 199 → 9%
11. Ј + ј = 2 + 63 = 65 → 3%
12. К + к = 1 + 59 = 60 → 3%
13. Л + л = 0 + 82 = 82 → 4%
14. Љ + љ = 0 + 10 = 10 → 0%
15. М + м = 2 + 95 = 97 → 4%
16. Н + н = 3 + 123 = 126 → 6%
17. Њ + њ = 0 + 23 = 23 → 1%
18. О + о = 3 + 290 = 293 → 13%
19. П + п = 1 + 55 = 56 → 2%
20. Р + р = 0 + 83 = 83 → 4%
21. С + с = 3 + 115 = 118 → 5%
22. Т + т = 2 + 100 = 102 → 4%
23. Ћ + ћ = 0 + 17 = 17 → 1%
24. У + у = 0 + 71 = 71 → 3%
25. Ф + ф = 0 + 5 = 5 → 0%
26. Х + х = 1 + 6 = 7 → 0%
27. Ц + ц = 0 + 9 = 9 → 0%
28. Ч + ч = 0 + 20 = 20 → 1%
29. Џ + џ = 0 + 0 = 0 → 0%
30. Ш + ш = 0 + 25 = 25 → 1%

У низу од 100 слова сличног текста, рецимо, слово „Д“ или „д“ појавиће се близу 5 пута на дужини од 100 слова текста, или тачније интензитет појављивања овог (великог или малог) слова азбуке, по једном слову текста, био је приближно λ = 0,0454 пута. То значи да очекујемо да се оно јавља након „сваких“ $\mu = 1/\lambda = 22$ слова текста. Уопште је:

$$F(\mu) = \int_0^{\mu} \rho(t)\,dt = 1 - e^{-\lambda\cdot\mu} = 1 - e^{-1} \approx 0{,}63212$$

вероватноћа да се слово појави у очекиваној дужини низа свог појављивања. ❏

Даљу расправу ове „вероватноће очекивања“, иначе само наговештене у датом примеру, видите у прилогу „Репрезентативни узорак“ (на енглеском). Узгред приметимо да би једнолика расподела вероватноћа слова у тексту указивала на њихову насумичност. Порука таквог текста била би попут излазне из црне кутије, канала толико оптерећеног шумом да улазне поруке на излазу „нема“.

Још једном, то би била потврда да губитак информације (дезинформацијом) настаје вишком (а не мањком) – због додатих сметњи самог канала преноса [53. Теорема 1]. Зато су насумичне поруке „без садржаја“, односно „неинтелигентне“ и имају униформне расподеле, јер је таква максималне информације [41. Теорема 1.*i*]. Уједно, када се појави „садржајност“, односно „смисленост“, као и организација система уопште видимо, на овај начин такође, да се смањује „количина опција“, или информација.

Међутим, знамо да теме о којима причамо могу бити повезане, да претходни делови текста могу иницирати следеће, када би се емисија словне информације понашала као да „има меморију“, те да за такве не можемо „отприлике“ утврдити имају ли експоненцијалну расподелу.

Ради комплетирања горњег примера ево и сортираних појављивања слова тог текста:

1. О 293	7. Д 103	13. У 71	19. Г 37	25. Љ 10
2. А 272	8. Т 102	14. Ј 65	20. Ш 25	26. Ц 9
3. И 199	9. М 97	15. К 60	21. Њ 23	27. Х 7
4. Е 192	10. Р 83	16. П 56	22. Ч 20	28. Ф 5
5. Н 126	11. Л 82	17. З 55	23. Ћ 17	29. Ђ 3
6. С 118	12. В 74	18. Б 50	24. Ж 14	30. Џ 0

Поделите наведени број са бројем свих слова текста (2268) и добићете фреквенцију, статистичку вероватноћу слова у датом тексту. Иако ту појава самог слова углавном следи експоненцијалну расподелу, верујемо, из ових бројева не видимо да то мора бити случај и са дистрибуцијом свих фреквенција. То питање „меморијске“ природе датог текста, да ли његове претходне реченице битно утичу на следеће, размотрићемо касније.

Пример 3. Из исте књиге „Приче о информацији“ узмимо следеће поглавље „1.2 Либерализам“ и избројмо велика и мала слова српске азбуке (компјутерским програмом). Од укупно 3938 знакова текста има 3217 великих или малих слова српске азбуке, сортираних у „топ листу“:

1. А 403 0,125	11. Д 116 0,036	21. Ш 30 0.009
2. И 326 0,101	12. Ј 104 0,032	22. Ћ 26 0.008
3. Е 292 0,091	13. У 104 0,032	23. Њ 25 0.008
4. О 292 0,091	14. К 93 0,029	24. Ж 22 0.007
5. Р 210 0,065	15. Л 93 0,029	25. Ц 21 0.007
6. Н 191 0,059	16. П 93 0,029	26. Ч 14 0.004
7. Т 142 0,044	17. Б 65 0,020	27. Љ 11 0.003
8. В 136 0,042	18. З 55 0,017	28. Ђ 8 0.002
9. С 135 0,042	19. Х 44 0,014	29. Ф 7 0.002
10. М 118 0,037	20. Г 41 0,013	30. Џ 0 0.000

Пласман у „топ листи“, слово, број појављивања међу 3217 слова и фреквенција. ❑

Програм за генерисање ових фреквенција у наставку садржи и иначе уграђене програме за линеарну регресију (Machine Learning - Linear Regression, Python) и цртање графа. То су кратки програми од којих ћу нешто навести касније. Један црта слику дистрибуције вероватноћа тачака лево (граф 1), а други (граф 2) израчунава линеарну регресију (оптималну праву линију средњих вредности) и приказује је поред истих тачака на слици десно.

Граф 1. и 2.

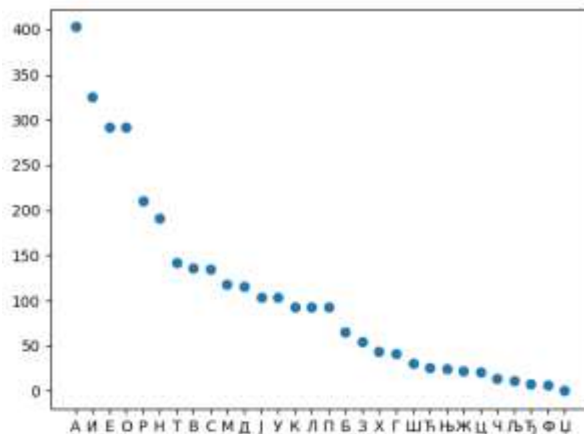
```
import matplotlib.pyplot as plt

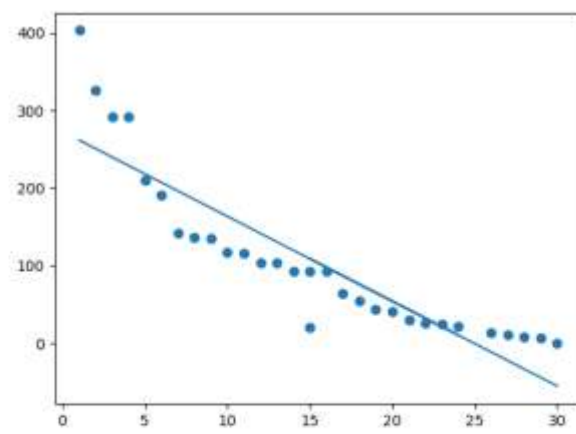
x = ['А', 'Б', 'В', 'Г', 'Д', 'Ђ', 'Е', 'Ж', 'З', 'И', 'Ј', 'К', 'Л',
'Љ', 'М', 'Н', 'Њ', 'О', 'П', 'Р', 'С', 'Т', 'Ћ', 'У', 'Ф', 'Х', 'Ц',
'Ч', 'Џ', 'Ш']

y =  [403, 326, 292, 292, 210, 191, 142, 136, 135, 118, 116, 104, 104,
93, 93, 93, 65, 55, 44, 41, 30, 26, 25, 22, 21, 14, 11, 8, 7, 0]

plt.scatter(x, y)

plt.show()
```






```
import matplotlib.pyplot as plt

from scipy import stats

x = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18,
19, 20, 21, 22, 23, 24, 15, 26, 27, 28, 29, 30]

y =  [403, 326, 292, 292, 210, 191, 142, 136, 135, 118, 116, 104, 104,
93, 93, 93, 65, 55, 44, 41, 30, 26, 25, 22, 21, 14, 11, 8, 7, 0]

slope, intercept, r, p, std_err = stats.linregress(x, y)

def myfunc(x):

  return slope * x + intercept

mymodel = list(map(myfunc, x))

plt.scatter(x, y)

plt.plot(x, mymodel)

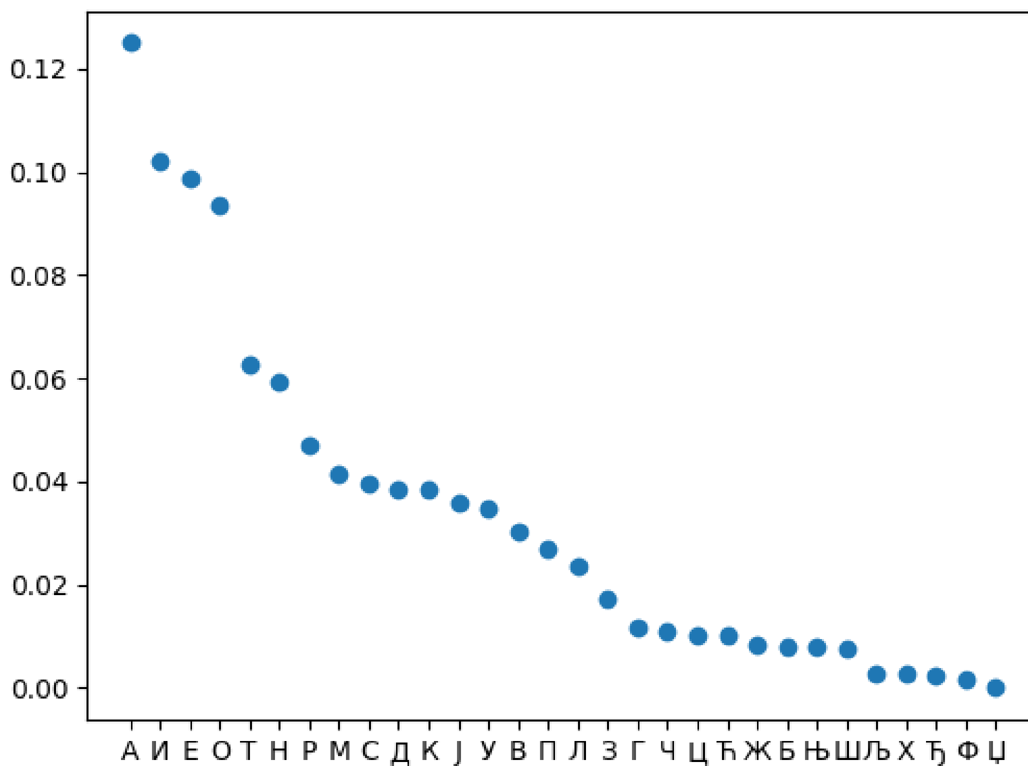
plt.show()
```

Крај примера графова. ❑

Пример 4. У наставку књиге „Приче о информацији“ је следеће поглавље „1.3 Истина или лаж“. Међу укупно 4108 тих знакова било је 3295 великих и малих слова азбуке, сортираних у:

1.	А	412	0.125	11.	К	127	0.039	21.	Ћ	33	0.010
2.	И	336	0.102	12.	Ј	118	0.036	22.	Ж	27	0.008
3.	Е	326	0.099	13.	У	115	0.035	23.	Б	26	0.008
4.	О	308	0.093	14.	В	100	0.030	24.	Њ	26	0.008
5.	Т	206	0.063	15.	П	89	0.027	25.	Ш	25	0.008
6.	Н	196	0.059	16.	Л	78	0.024	26.	Љ	9	0.003
7.	Р	155	0.047	17.	З	57	0.017	27.	Х	9	0.003
8.	М	137	0.042	18.	Г	38	0.012	28.	Ђ	8	0.002
9.	С	131	0.040	19.	Ч	36	0.011	29.	Ф	6	0.002
10.	Д	127	0.039	20.	Ц	34	0.010	30.	Џ	0	0.000

За разлику од претходних примера на следећем графу је дистрибуција вероватноћа ових слова.



Подсећам, 412 је број појављивања слова 'А' у датом тексту. Он подељен бројем 3295 свих слова је вероватноћа тог слова, $\Pr(\mathrm{A}) = 0.125$, која је последњи број колоне. ❑

У сва три примера водеће фреквенције имају неки самогласници који су у групама 1-3, а за њима следе све ређа слова азбуке. Тешко је напамет рећи да ли је ова дистрибуција (свих слова азбуке) експоненцијална, што би значило независност следећих делова текста од претходних.

Унапред дато слово азбуке у тексту може се појављивати у експоненцијалној расподели, а да за дистрибуцију свих слова азбуке то не мора бити. У случају насумичног текста, дистрибуција слова биће униформна (максималне вероватноће на интервалу те азбуке), али ће у случају смислених порука имати чешћа и ређа слова. Зато ће „смислен“ текст (са поруком) имати мању информацију од „бесмисленог“ (насумичног)!

Ово запажање је потврда, иначе кажу ми „апсурдног“ становишта моје „теорије информације“, да саопштавањем порука шаљемо мање информације. То је зато што је шум (сметња) додатак који не брише оригинале већ их само замагљује, дезинформише нас. Доследно у мојој наводној теорији, организацијом могућности смањујемо њихову „количину опција“ (информацију), или слично, са организацијом друштва смањују се слободе (просечан обим избора) индивидуа. То су заправо и неки од главних садржаја расправљаних у овој скрипти, попут привлачности минимализма.

64.3. Раздвојивост

Ерланг је основна јединица телекомуникационог саобраћаја коју користе дизајнери говорних система – названа по данском математичару, статистичару и инжењеру, који је изумео области саобраћајног инжењерства и теорије чекања. Изнети промет у ерлангима је просечан број позива измерен у датом периоду (често један сат), док је понуђени саобраћај онај који би се преносио ако би сви покушаји били успешни. Овде ћемо нагласити само један од аспеката Ерлангових процеса, њихову сепарабилност (раздвојивост).

Теорема 2. (Ерлангов процес) Претпоставимо да су испуњени сви услови за експоненцијални процес и да важи сепарабилност, тј. да је појава два или више различитих сигнала у истом тренутку (скоро сигурно) немогућ догађај. Тада је вероватноћа појаве тачно $j = 0,1,\dots,s$ сигнала у интервалу времена $t > 0$ једнака

$$F_j(t) = \frac{(\lambda t)^j}{j!} \cdot e^{-\lambda t}$$

где је параметар $\lambda > 0$.

Доказ: Према формули потпуне вероватноће је

$$F_j(t + \Delta t) = \sum_{i=0}^{j} F_i(t) \cdot F_{j-1}(\Delta t).$$

Наиме, пријем j сигнала у интервалу $t + \Delta t$ може се раставити на $j + 1$ независан догађај:

1. 0 сигнала у t, а j сигнала у Δt;
2. 1 сигнал у t, а $j - 1$ сигнал у Δt;

…

j+1. j сигнала у t, а 0 сигнала у Δt.

Узмемо ли довољно мало $\Delta t \to 0$, тада претпоставка сепарабилности искључује ситуације од 1 до j. Зато је

$$F_j(t+\Delta t) = F_j(t)\cdot F_0(\Delta t) + F_{j-1}(t)\cdot F_1(\Delta t).$$

Из истог разлога ($\Delta t \to 0$) се у интервалу Δt може појавити један или ниједан сигнал, тј.

$$F_0(\Delta t) + F_1(\Delta t) = 1.$$

Отуда је

$$F_j(t+\Delta t) - F_j(t) = \left[F_{j-1}(t) - F_j(t)\right]\cdot F_1(\Delta t).$$

Овде је $F_0(\Delta t)$ вероватноћа да се неће појавити нити један од сигнала у интервалу Δt под претпоставком важења експоненцијалног процеса, те је:

$$F_1(\Delta t) = 1 - F_0(\Delta t) = 1 - e^{-\lambda\Delta t} = \lambda\cdot\Delta t + o(\Delta t).$$

Преласком на лимес, када $\Delta t \to 0$, добијамо једначине

$$\frac{dF_j(t)}{dt} = -\lambda\cdot F_j(t) + \lambda\cdot F_{j-1}(t).$$

са почетним условима $F_0(0) = 1$, $F_j(0) = 0$ за $j \geq 1$. Лако се решавају сменом $F_j(t) = e^{-\lambda t} f_j(t)$, где су $f_j(t)$ непознате функције. Добијамо

$$\frac{df_j(t)}{dt} = \lambda\cdot f_{j-1}(t)$$

где је $f_0(t) = 1$. Рекурзије дају:

$$f_j(t) = \frac{(\lambda\cdot t)^j}{j!},\quad \text{тј.}\quad F_j(t) = \frac{(\lambda t)^j}{j!}\cdot e^{-\lambda t}.$$

Тиме је теорема доказана. ■

Погледајмо како се иста теорема (Ерлангов процес) налази из Колмогоровљевих диференцијалних једначина. Због сепарабилности, за мали интервал Δt у којем се може одвијати највише по једна трансформација, $\widehat{K}(\Delta t)$, могућа је само једна промена стања сисгтема. Према томе, за дато i само је један од коефицијената k_{ij} ($i \neq j$) различит од нуле. Ако је F_i вероватноћа реализације i-тог стања (стигло је i од s сигнала), онда систем за време Δt може прећи највише у $i+1$ стање.

Због независности је густина вероватноће међусобних прелаза константна, па за свако i је:

$$\lambda_{ii} = -\lambda,\quad \lambda_{i,i+1} = \lambda$$

док су сви остали коефицијенти нула. То је довољно да се формира систем диференцијалних једначина

$$\frac{dF_i(t)}{dt} = -\lambda\cdot F_i(t) + \lambda\cdot F_{i-1}(t)$$

чије решење је Ерлангов процес.

Ерлангов процес се може добити и из тзв. Бернулијеве шеме која посматра само један случајни догађај π који се дешава са вероватноћом p, или изостаје са вероватноћом $q = 1 - p$. Ако опит понављамо n пута, онда Бернулијева вероватноћа да ће се догађај π десити тачно $j = 0,1, \dots, n$ пута износи

$$P_n(j) = \binom{n}{j} p^j q^{n-j}.$$

Ако $n \cdot P_n(j) \to a$, када је j фиксирано и $n \to \infty$, тада

$$P_n(j) \to \frac{a^j}{j!} \cdot e^{-a}.$$

Детаљна објашњења и доказе можете видети у мојој књизи Квантна Механика, 1.1.7 Расподеле вероватноћа. Посебно само о сепарабилности касније ћемо се позабавити мало детаљније.

Теорема 3. (Пуасонов став) Ако је

$$f_n(k) = \binom{n}{k} p_n^k (1 - p_n)^{n-k}$$

и ако низ бројева $p_n = a_n \to a$ када $n \to \infty$, тада је

$$\lim_{н \to \infty} f_n(k) = \frac{a^k}{k!} \cdot e^{-a}.$$

Доказ:

$$f_n(k) = \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{n!}{k!\,(n-k)!} \cdot \left(\frac{a_n}{n}\right)^k \cdot \left(1 - \frac{a_n}{n}\right)^{n-k} =$$

$$= \frac{a_n^k}{k!} \cdot \left(1 - \frac{a_n}{n}\right)^n \cdot \frac{\left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right)}{\left(1 - \frac{a_n}{n}\right)^k}.$$

Када $n \to \infty$ трећи фактор, велики разломак, тежи 1, средњи тежи e^{-a}, а први $a^k/k!$. Тиме је став доказан. ■

Када у Бернулијевој шеми меримо n пута појаву догађаја π у току времена t, а када за исто време $n \to \infty$ добићемо $n \cdot t \cdot p_n \to a$, односно $n \cdot p_n \to \lambda$, док је $a = \lambda t$. Ова је апроксимација нарочито погодна за $n > 50$ и $np < 10$. Из Пуасонове добијамо Ерлангову, то брже што је a мање.

Хаотично кретање честица, познато као Брауново кретање, има ову особину. Ако се апсиса једне честице помера за x улево или удесно према Бернулијевој шеми, онда ће кретање много честица описивати гранични процес Пуасоновог става.

Са друге стране имамо услове Ерланговог процеса:

i. кретање честице не зависи од избора почетног тренутка;
ii. честице у различитим областима се крећу независно;
iii. немамо две или више честица у истој тачки.

Према томе је јасан смисао вероватноће

$$F_\lambda(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \quad \lambda > 0$$

где је $\lambda = np$ очекивање, p вероватноћа сваког појединог исхода у низу од n опита, при чему је жељени број реализација тачно k. Овако формулисан Пуасонов процес лако се проверава и добро апроксимира Бернулијеву расподелу за $\lambda < 10$ и $n > 50$.

Пример 5. У претходном, 4. примеру, вероватноћа појављивања слова 'Ф' је $p = 0.002$. Тражимо вероватноће да у случајно бираном низу од 100 слова тог текста:

1) буде тачно 5 слова 'Ф';
2) број слова 'Ф' није већи од 5.

Из вероватноће да је насумично слово управо тражено, $\Pr(\Phi) = p = 0.002$, следи вероватноћа да такво (поједино) слово није тражено, $\Pr(\neg\Phi) = 1 - p = q = 0.998$. Како низ од 100 слова можемо схватити као независно понављање једног слова 100 пута, то је према Бернулијевој шеми:

1) $f_{100}(k = 5) = \Pr\{k = 5\} = \binom{100}{5} \times 0.002^5 \times 0.998^{95} \approx 0.000002;$
2) $f_{100}(k \leq 6) = \Pr\{k \leq 5\} = \Pr\{k = 0\} + \Pr\{k = 1\} + \cdots + \Pr\{k = 5\} =$

$$\approx 0.818567 + 0.164041 + 0.016273 + 0.001065 + 0.000052 + 0.000002 \approx 0.99.$$

Међутим, можемо применити и Пуасонову апроксимацију, јер $\lambda = np = 100 \times 0.002 = 0.2 < 10$ и $n = 100 > 50$. Тако израчунавамо:

1) $\Pr\{k = 5\} = \frac{\lambda^k}{k!} \cdot e^{-\lambda} = \frac{0.2^5}{5!} \cdot e^{-0.2} \approx 0.000002;$
2) $\Pr\{k \leq 5\} = \sum_{k=0}^{5} \frac{0.2^k}{k!} e^{-0.2} = e^{-0.2}\left(\frac{1}{1} + \frac{0.2}{1} + \frac{0.04}{2} + \frac{0.008}{6} + \frac{0.0016}{24} + \frac{0.00032}{120}\right) \approx 0.999.$

Резултати су једнаки у првих шест децимала, а бивали би још више уједначени за веће n и мање p. Случајни исечак датог текста: „настојећи да ту нешто поправимо у незнању да смо заправо жртве шкртарења природе у давању информације. Грци су приметили“ – као што се види, има 101 слово и једно појављивање слова „ф“. Највероватније ($p_0 \approx 0.82$) да у низу од око 100 слова, слова 'ф' ни неће бити. ❑

Очекивање и варијанса Пуасонове расподеле су $\mathrm{E}(X) = \mathrm{Var}(X) = \lambda$, у 5. примеру $\lambda = 0.02$, што се види да је такође тачно кроз 2, 3. и 4. пример, јер се вероватноће слова 'ф' тамо појављују редом:

$$\frac{5}{2268} \approx 0.002205, \quad \frac{7}{3217} = 0.002176, \quad \frac{6}{3295} = 0.001821.$$

Варијације су у оквиру израчунате варијансе, а слично је и са Бернулијевом шемом. То је уједно и врста „експерименталне“ потврде да се поједина слова у тексту држе експоненцијалне расподеле, конкретно Пуасонове. Такође, да је појављивање унапред датог слова у низу слова текста лишено меморије.

65. Регресија

Вратимо се на примере графова [64. Граф 2] дистрибуције слова у тексту и праве линије која их најбоље представља. Узмимо сада следећи прилог (1.4 Феминизација) из исте књиге „Приче о информацији“, пребројмо све знакове новог текста (3850) и посебно уочимо и пребројмо сва ћирилична велика и мала слова (3150). Резултат је:

А + а = 3 + 373 = 376
Б + б = 1 + 39 = 40
В + в = 1 + 102 = 103
Г + г = 1 + 45 = 46
Д + д = 6 + 93 = 99
Ђ + ђ = 0 + 6 = 6
Е + е = 1 + 300 = 301
Ж + ж = 0 + 14 = 14
З + з = 4 + 66 = 70
И + и = 5 + 320 = 325
Ј + ј = 0 + 110 = 110
К + к = 3 + 94 = 97
Л + л = 1 + 73 = 74
Љ + љ = 0 + 14 = 14
М + м = 0 + 121 = 121
Н + н = 3 + 205 = 208
Њ + њ = 0 + 35 = 35
О + о = 2 + 284 = 286
П + п = 4 + 90 = 94
Р + р = 0 + 161 = 161
С + с = 3 + 155 = 158
Т + т = 2 + 136 = 138
Ћ + ћ = 0 + 21 = 21
У + у = 2 + 129 = 131
Ф + ф = 1 + 15 = 16
Х + х = 0 + 17 = 17
Ц + ц = 0 + 39 = 39
Ч + ч = 0 + 23 = 23
Џ + џ = 0 + 0 = 0
Ш + ш = 0 + 27 = 27

Након сортирања по броју појављивања слова, добијамо њихов „пласман“:

1. А 376 0.119
2. И 325 0.103
3. Е 301 0.096
4 О 286 0.091
5. Н 208 0.066
6. Р 161 0.051
7. С 158 0.050
8. Т 138 0.044
9. У 131 0.042
10. М 121 0.038
11. Ј 110 0.035
12. В 103 0.033
13. Д 99 0.031
14. К 97 0.031
15. П 94 0.030
16. Л 74 0.023
17. З 70 0.022
18. Г 46 0.015
19. Б 40 0.013
20. Ц 39 0.012
21. Њ 35 0.011
22. Ш 27 0.009
23. Ч 23 0.007
24. Ћ 21 0.007
25. Х 17 0.005
26. Ф 16 0.005
27. Ж 14 0.004
28. Љ 14 0.004
29. Ђ 6 0.002
30. Џ 0 0.000

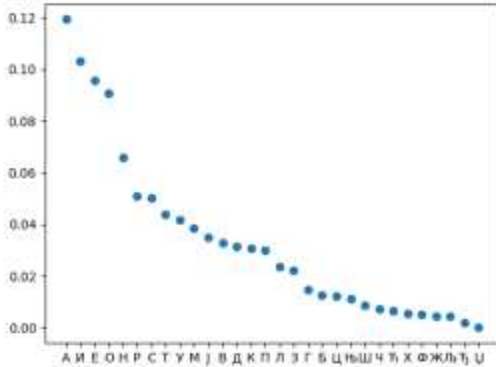
У последњој колони су фреквенције, тј. „статистичке вероватноће“. То су бројеви претходне колоне дељени константним бројем свих слова (3150).

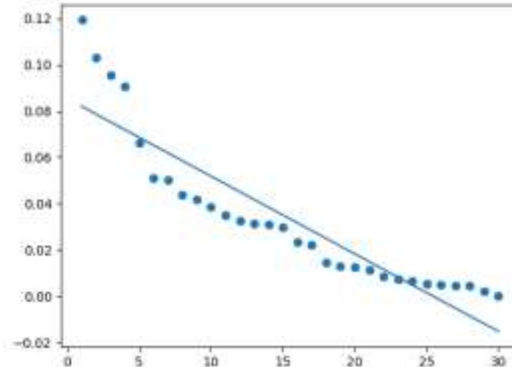
65.1. Праволинијска

На следеће две слике су графови дистрибуције горњих 30 великих и малих слова српске ћирилице. На левој, дуж апсцисе, доле потписано словима како би се могао видети њихов поредак, иначе из претходног пласмана „топ листе“. Десно је иста дистрибуција са графом праве линије (линеарне регресије) која их „најбоље“ апроксимира.

Тачке дистрибуције дате су паровима координата (x_k, y_k), за $k = 1,2, \dots, s$, где је у нашем случају број слова $s = 30$. На пример $\text{Т}(8, 0.044)$. Права линија регресије је $f_k = f(x_k) = \alpha_1 x_k + \alpha_2$, са α_1 и α_2 константним параметрима које ћемо сада и израчунати. Низ x_k су, као што видимо, првих 30

природних бројева, а константе (α_1 и α_2), тражимо методом најмањих квадрата, тј. тако да збир свих 30 квадрата разлика ордината датих тачака (y_k) и ордината регресионе линије (f_k) буде што је могуће мањи.





Другим речима, тражимо такве α_1 и α_2 да функција

$$g(\alpha_1, \alpha_2) = (f_1 - y_1)^2 + (f_2 - y_2)^2 + \cdots + (f_s - y_s)^2 =$$

$$= (\alpha_1 x_1 + \alpha_2 - y_1)^2 + (\alpha_1 x_2 + \alpha_2 - y_2)^2 + \cdots + (\alpha_1 x_s + \alpha_2 - y_s)^2$$

$$= \sum_{k=1}^{s} (\alpha_1 x_k + \alpha_2 - y_k)^2$$

има најмању вредност:

$$\frac{\partial g}{\partial \alpha_1} = 0, \quad \frac{\partial g}{\partial \alpha_2} = 0,$$

$$\sum_{k=1}^{s} 2(\alpha_1 x_k + \alpha_2 - y_k) x_k = 0, \quad \sum_{k=1}^{s} 2(\alpha_1 x_k + \alpha_2 - y_k) = 0,$$

$$\alpha_1 \sum_{k=1}^{s} x_k^2 + \alpha_2 \sum_{k=1}^{s} x_k = \sum_{k=1}^{s} x_k y_k, \quad \alpha_1 \sum_{k=1}^{s} x_k + \alpha_2 s = \sum_{k=1}^{s} y_k$$

$$\alpha_1 A_2 + \alpha_2 A_1 = B_1, \quad \alpha_1 A_1 + \alpha_2 s = B_0$$

где су бројеви:

$$A_2 = \sum_{k=1}^{s} x_k^2 = 9455, \quad A_1 = \sum_{k=1}^{s} x_k = 465, \quad B_1 = \sum_{k=1}^{s} x_k y_k = 7.958095 \quad B_0 = \sum_{k=1}^{s} y_k = 1,$$

познати ($s = 30$). Систем линеарних једначина:

$$\begin{cases} \alpha_1 A_2 + \alpha_2 A_1 = B_1 \\ \alpha_1 A_1 + \alpha_2 s = B_0 \end{cases}$$

има решење:

$$\begin{cases} \alpha_1 = \dfrac{sB_1 - A_1B_0}{sA_2 - A_1^2} = \dfrac{s \cdot (\sum_{k=1}^{s} x_k y_k) - (\sum_{k=1}^{s} x_k)}{s \cdot (\sum_{k=1}^{s} x_k^2) - (\sum_{k=1}^{s} x_k)^2} \\ \alpha_2 = \dfrac{A_2B_0 - A_1B_1}{sA_2 - A_1^2} = \dfrac{(\sum_{k=1}^{s} x_k^2) - (\sum_{k=1}^{s} x_k) \cdot (\sum_{k=1}^{s} x_k y_k)}{s \cdot (\sum_{k=1}^{s} x_k^2) - (\sum_{k=1}^{s} x_k)^2} \end{cases}$$

$$\begin{cases} \alpha_1 = \dfrac{30 \cdot 7.958095 - 465}{30 \cdot 9455 - 465^2} = \dfrac{-226.25715}{67425} = -0.0033556862 \\ \alpha_2 = \dfrac{9455 - 465 \cdot 7.958095}{30 \cdot 9455 - 465^2} = \dfrac{5754.485825}{67425} = 0.0853464696 \end{cases}$$

Добили смо граф приближно праве линије, једначине $f(x) = -0.003x + 0.085$. Коефицијенти су јој мали бројеви због размере координатног система. Ординате су делови јединице (вероватноће) а апсцисе су природни бројеви (30 пута већи), а затим њу равнају и учешћа многих ретких слова.

Средње вредности (очекивања) датих и израчунатих тачака су:

$$\mu_y = \sum_{k=1}^{s} x_k y_k \quad \mu_{\text{lin}} = \sum_{k=1}^{s} x_k f_k$$

$$\mu_y = B_1 \quad \mu_{\text{lin}} = \alpha_1 A_2 + \alpha_2 A_1$$

дакле једнаке су ($\alpha_1 A_2 + \alpha_2 A_1 = B_1$), овде $\mu = \mu_y = \mu_{\text{lin}} = 7.9580952380952$, са разликеама (39 и 41) тек након 13. децимале, а и то због грешака заокруживања. Исте су и варијансе:

$$\sigma_y^2 = \frac{1}{s}\sum_{k=1}^{s}\left(x_k - \mu_y\right)^2 = \frac{1}{s}\sum_{k=1}^{s}(x_k - \mu_{\text{lin}})^2 = \sigma_{\text{lin}}^2 = \sigma^2 = 131.796994,$$

па су им исте и дисперзије $\sigma_y = \sigma_{\text{lin}} = \sqrt{\sigma^2} = 11.480287$, што је разумљиво и интуитивно, јер су то мере средње удаљености иначе фиксних тачака апсцисе ($x_k = k$) од једне њене фиксне тачке (μ).

Како је број алфа1 негативан, а број алфа2 је релативно мали ($\alpha_1 < 0$ и $\alpha_2 \approx 0$), то редни бројеви на крају листе дају негативне ординате ($f_k = \alpha_1 x_k + \alpha_2$), наводне вероватноће слова (са редним бројем $x_k > 25$). Оне су, дакле, нетачне и чине додатни проблем употребе ове регресије.

У следећој табели наводимо претходну „топ листу“ упоредо са израчунатим вероватноћама, нове према формули $f_k = \alpha_1 x_k + \alpha_2$ са коефицијентима $\alpha_1 = -0.003356$, $\alpha_2 = 0.085346$. Последње негативне израчунате „вероватноће“ такође су у табели, иако таквих стварно нема:

1.	0.119365	0.08199	11.	0.034921	0.04843	21.	0.011111	0.01487
2.	0.103175	0.078634	12.	0.032698	0.045074	22.	0.008571	0.011514
3.	0.095556	0.075278	13.	0.031429	0.041718	23.	0.007302	0.008158
4.	0.090794	0.071922	14.	0.030794	0.038362	24.	0.006667	0.004802
5.	0.066032	0.068566	15.	0.029841	0.035006	25.	0.005397	0.001446
6.	0.051111	0.06521	16.	0.023492	0.03165	26.	0.005079	-0.00191
7.	0.050159	0.061854	17.	0.022222	0.028294	27.	0.004444	-0.005266
8.	0.04381	0.058498	18.	0.014603	0.024938	28.	0.004444	-0.008622
9.	0.041587	0.055142	19.	0.012698	0.021582	29.	0.001905	-0.011978
10.	0.038413	0.051786	20.	0.012381	0.018226	30.	0.000000	-0.015334

Зато ни $\sum f_k$ није 1, а требало би бити. Сличан је проблем са Шеноновом информацијом, не датих, него регресијом добијених вредности. Ипак наводим овај рачун:

$$S_y = -\sum_{y_k>0} y_k \log_2 y_k = 4.292664 \quad S_{\text{lin}} = -\sum_{f_k>0} f_k \log_2 f_k = 4.486895$$

Узимам само позитивне вероватноће (свих 15 децимала, компјутером). Испада да је информација „расутих“ S_y (ордината) ипак мања од „уједначене“ информације S_{lin} (линије). То је прихватљиво, јер „хоризонталније“ вероватноће, равноправнијих исхода, имају већу информацију.

За крај дефинишимо и запамтимо „дистанцу“ између датих и овако израчунатих вредности:

$$d_{\text{lin}} = d(f_{\text{lin}}, y) = \sqrt{(f_1 - y_1)^2 + \cdots (f_s - y_s)^2} = 0.072938$$

Тај број d_{lin} утолико је мањи што је боље налегање линије на тачке дате расподеле. За разлику од дисперзије $\sigma \approx 11.480$, мере просечне удаљености свих апсциса од фиксиране, ово $d_{\text{lin}} \approx 0.073$ је метрика, растојање између низова (врхова вектора).

65.2 Параболна

Узмио опет исте претходне тачке дистрибуције означене паровима координата (x_k, y_k) индекса $k = 1,2, \dots, s$. Број слова је $s = 30$ и, на пример $\mathrm{E}(3,\ 0.096)$. Линија регресије биће сада квадратна функција $f_{\text{par}} = f(x) = \beta_1 x^2 + \beta_2 x + \beta_3$, константних параметара β_1, β_2 и β_3 које требамо израчунати.

Првих $s = 30$ природних бројева x_k дефинишу ординате $f_k = f(x_k) = \beta_1 x_k^2 + \beta_2 x_k + \beta_3$ параболе f_{par} коју тражимо – тако да су разлике $|f_k - y_k|$ што је могуће мање. Користимо опет исту методу „најмањих квадрата“:

$$g(\beta_1, \beta_2, \beta_3) = \sum_{k=1}^{s} (f_k - y_k)^2 = \sum_{k=1}^{s} \left(\beta_1 x_k^2 + \beta_2 x_k + \beta_3 - y_k\right)^2$$

Третирајући β_1, β_2 и β_3 као непознате променљиве, са низовима x_k и y_k датих константи:

$$\frac{\partial g}{\partial \beta_1} = 0, \quad \frac{\partial g}{\partial \beta_2} = 0, \quad \frac{\partial g}{\partial \beta_3} = 0,$$

$$\beta_1 A_4 + \beta_2 A_3 + \beta_3 A_2 = B_2, \quad \beta_1 A_3 + \beta_2 A_2 + \beta_3 A_1 = B_1, \quad \beta_1 A_2 + \beta_2 A_1 + s\beta_3 = B_0,$$

где је:

$$A_4 = \sum_{k=1}^{s} x_k^4 = 5273999, \qquad A_3 = \sum_{k=1}^{s} x_k^3 = 216225,$$

$$A_2 = \sum_{k=1}^{s} x_k^2 = 9455, \qquad A_1 = \sum_{k=1}^{s} x_k = 465,$$

$$B_0 = 1, \qquad B_1 = \sum_{k=1}^{s} x_k y_k = 7.958095, \qquad B_2 = \sum_{k=1}^{s} x_k^2 y_k = 104.390476.$$

Ово су коефицијенти система линеарних једначина по непознатима β_1, β_2 и β_3:

$$\begin{cases} \beta_1 A_4 + \beta_2 A_3 + \beta_3 A_2 = B_2 \\ \beta_1 A_3 + \beta_2 A_2 + \beta_3 A_1 = B_1 \\ \beta_1 A_2 + \beta_2 A_1 + \beta_3 s = B_0 \end{cases}$$

Непознате тражимо Крамеровом, методом детерминанти ($s = 30$):

$$\Delta = \begin{vmatrix} A_4 & A_3 & A_2 \\ A_3 & A_2 & A_1 \\ A_2 & A_1 & s \end{vmatrix} = 9051851200.000034, \quad \Delta_a = \begin{vmatrix} B_2 & A_3 & A_2 \\ B_1 & A_2 & A_1 \\ B_0 & A_1 & s \end{vmatrix} = 1552316.1428571267$$

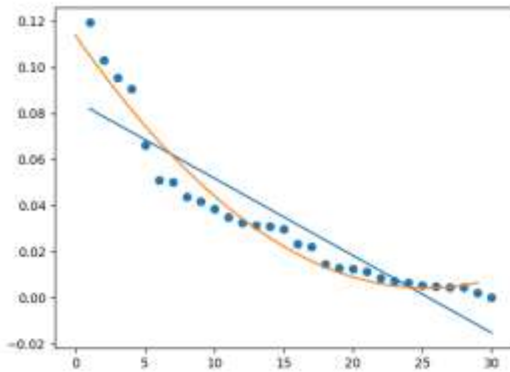
$$\Delta_b = \begin{vmatrix} A_4 & B_2 & A_2 \\ A_3 & B_1 & A_1 \\ A_2 & B_0 & s \end{vmatrix} = -78496972.6952381, \quad \Delta_c = \begin{vmatrix} A_4 & A_3 & B_2 \\ A_3 & A_2 & B_1 \\ A_2 & A_1 & B_0 \end{vmatrix} = 1029193145.7523854$$

$$\beta_1 = \frac{\Delta_a}{\Delta} = 0.000171, \quad \beta_2 = \frac{\Delta_b}{\Delta} = -0.008672, \quad \beta_3 = \frac{\Delta_c}{\Delta} = 0.113700.$$

Парабола регресије, квадратна функција $f(x) = \beta_1 x^2 + \beta_2 x + \beta_3$, има минимум $f_0 = -3 \times 10^{-6}$ у тачки $x_0 = -\frac{\beta_2}{2\beta_1} = 25.283823$, што значи да њене вредности након 25. слова расту. То се види у упоредној табели „топ листе“:

1.	0.119365	0.105199	11.	0.034921	0.039059	21.	0.011111	0.007217
2.	0.103175	0.097042	12.	0.032698	0.034331	22.	0.008571	0.005919
3.	0.095556	0.089227	13.	0.031429	0.029947	23.	0.007302	0.004965
4.	0.090794	0.081756	14.	0.030794	0.025905	24.	0.006667	0.004353
5.	0.066032	0.074627	15.	0.029841	0.022206	25.	0.005397	0.004084
6.	0.051111	0.067842	16.	0.023492	0.018851	26.	0.005079	0.004158
7.	0.050159	0.061399	17.	0.022222	0.015838	27.	0.004444	0.004575
8.	0.043810	0.055300	18.	0.014603	0.013168	28.	0.004444	0.005335
9.	0.041587	0.049543	19.	0.012698	0.010842	29.	0.001905	0.006438
10.	0.038413	0.044130	20.	0.012381	0.008858	30.	0.000000	0.007884

У првој колони табеле је „пласман“ слова, у другој је дата фреквенција, у трећој израчуната. Види



се да након 25. позиције дате вероватноће и даље опадају, док параболна расте. То значи да ни квадратну, ни праволинијску регресију, сваку због њених посебности, не треба сувише развлачити у применама. Међутим, нови граф, квадратне регресије, значајно боље лежи по датим тачкама, као што показује слика лево.

Плаве тачке представљају дате вероватноће слова ћирилице, плава линија праволинијску регресију, а жута је регресија параболом. Иста су очекивање нове регресије претходним, јер:

$$\mu_{\text{par}} = \sum_{k=1}^{s} x_k f_k = \sum_{k=1}^{s} x_k\left(\beta_1 x_k^2 + \beta_2 x_k + \beta_3\right) =$$

$$\beta_1 \sum_{k=1}^{s} x_k^3 + \beta_2 \sum_{k=1}^{s} x_k^2 + \beta_3 \sum_{k=1}^{s} x_k = \beta_1 A_3 + \beta_2 A_2 + \beta_3 A_1 = B_1 = \mu_y.$$

Дакле, средње вредности (математичка очекивања) параболне, праволинијске регресије и дате дистрибуције једнака су:

$$\mu_{\text{par}} = \mu_{\text{lin}} = \mu_y = \mu = B_1 = 7.958095.$$

Варијанса и дисперзија квадратне регресије износе:

$$\sigma_{\text{par}}^2 = \frac{1}{s}\sum_{k=1}^{s}(f_k - \mu)^2 = 62.802826, \quad \sigma_{\text{par}} = \sqrt{\sigma_{\text{par}}^2} = 7.924823$$

Па су три дисперзије (дате, линијске и параболне регресије), корени варијанси:

$$\sigma_y = 7.924826, \quad \sigma_{\text{lin}} = 7.924815, \quad \sigma_{\text{par}} = 7.924823.$$

у поретку су: $\sigma_{\text{lin}} < \sigma_{\text{par}} < \sigma_y$, али разлике су занемарљиве. Шенонова информација је

$$S_{\text{par}} = -\sum_{f_k > 0} f_k \log_2 f_k = 4.295455$$

Три (једнако дефинисане) Шенонове информације ових расподела:

$$S_y = 4.292664, \quad S_{\text{lin}} = 4.486895, \quad S_{\text{par}} = 4.295455,$$

су у поретку: $S_y < S_{\text{par}} < S_{\text{lin}}$. Дистанце вектора линеарне и квадратне регресије од дате су:

$$d_{\text{lin}} = d(f_{\text{lin}}, y) = 0.072938, \quad d_{\text{par}} = d(f_{\text{par}}, y) = 0.037036.$$

Највише су уједначене вероватноће линеарне регресије и таква има највећу информацију, јер јој је неизвесност, а онда и емитована информација, максимална. Информација равноправнијих шанси већа је. Код дате расподеле разлике вероватноћа највеће су и она од ових израчунатих има мању информацију.

Ово запажање нам може изгледати „чудно“, али приметимо да је доследно. Садржајан, смислен, или паметнији говор је усмеренији, организованији, он је редукован од неизвесности. Слично је са повећавањем информације канала преноса порука док се додатно оптерећује сметњама (шумом). Такође, боље се уклапајући у боље организовано друштво губимо неке личне слободе (количине опција, тј. информације). Тежећи сигурности одричемо се неизвесности, уклапајући се у поредак одустајемо од „непотребних“ опција, од ризика. Такође, то значи да расту неједнакости могућих наших избора (знамо шта нам је чинити), смањују се „штетне“ дилеме и обим наших слобода је заправо мањи. Тада је мања Шенонова информација наших одлука.

Познато нам је да задаци тражења експоненцијалних регресија методама најмањих квадрата нису лако решиви, пре свега због незгодних трансцедентних једначина које нам тада стоје на путу. То се делом види из примера који следи, а затим ћемо савладати и тај проблем, на другачији начин.

65.3. Експоненцијална

Нека је дата иста дистрибуција слова, са координатама тачака $(x_k = k, y_k)$ и потражимо најближу њој експоненцијалну расподелу. То можемо покушати на претходни начин, тј. методом најмањих квадрата, као што је приказано у следећим примеру.

Пример 1. функција $f_\lambda = f(x) = e^{-\lambda x}$ има коефицијент $\lambda > 0$ који тражимо методом најмањих квадрата. То значи, зависно од параметра λ, наћи минимум функције:

$$g(\lambda) = \sum_{k=1}^{s} (f_k - y_k)^2 = \sum_{k=1}^{s} \left(e^{-\lambda k} - y_k\right)^2.$$

За стационарну тачку, налазимо извод g по λ, тј. $\frac{\partial g}{\partial \lambda} = 0$, односно $\sum_{k=1}^{s} -k\left(e^{-\lambda k} - y_k\right) = 0$. Следи да су очекиване вредности дате и ове расподеле једнаке ($\mu_y = \mu_e = \mu$), те:

$$\mu = \sum_{k=1}^{s} k y_k = \sum_{k=1}^{s} k e^{-\lambda k} = -\sum_{k=1}^{s} \frac{\partial}{\partial \lambda}\left(e^{-\lambda k}\right) = -\frac{\partial}{\partial \lambda}\left(\sum_{k=1}^{s} e^{-\lambda k}\right) =$$

$$= -\frac{\partial}{\partial \lambda}\left(e^{-\lambda} \cdot \frac{1 - e^{-s\lambda}}{1 - e^{-\lambda}}\right) = e^{-\lambda} \cdot \frac{1 - (s+1)e^{-s\lambda} + s e^{-(s+1)\lambda}}{(1 - e^{-\lambda})^2}.$$

Међутим, већ за једноцифрене бројеве s, те за λ не много веће од 1, биће

$$s e^{-(s+1)\lambda} \approx (s+1) e^{-s\lambda}$$

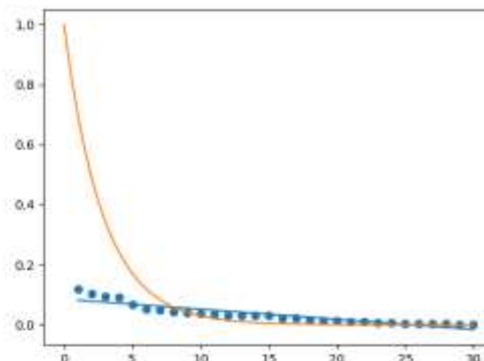
па за $s \to \infty$ са добром тачношћу имамо процену

$$\mu = e^{-\lambda} / \left(1 - e^{-\lambda}\right)^2$$

У датом примеру је $\mu = 7.958$, а томе одговара $\lambda = 0.353$. Видимо тај експоненцијални граф

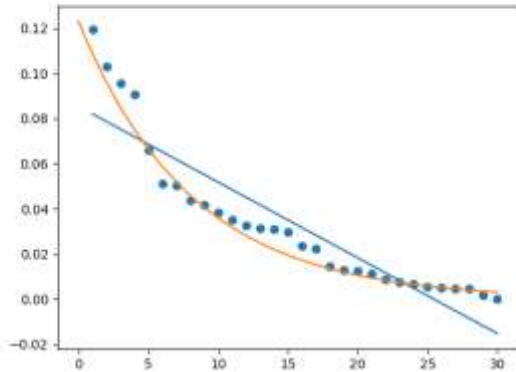
$$f(x) = e^{-0,353x}$$

жуто бојен на слици десно, упоредо са плавом правом линијом од пре, као и дате плаве тачке фреквенције (вероватноће) слова. ❑



Веома много вероватноће (жутог графа) троши се на велома мало слова, па околина очекиваног (око 8) има заравњен, скоро хоризонталан граф, баш у „епицентру“. Затим, површина испод овог графа није јединична, те он не представља расподелу вероватноћа.

Други начин тражења најближе експоненцијалне расподеле датој је одмах претпоставком да две расподеле имају исто очекивање, тј. средњу вредност најзначајнијег места апсцисе. Посебно овде приближно $x = \mu = 7.958$ даје $\gamma = 1/\mu = 0.126$, а потом „најбољим“ просто проглашавамо граф $f_\gamma(x) = \gamma e^{-\gamma x}$, доследно разматрањима из поднаслова „64.1. Независност“.



Граф ове експоненцијалне функције регресије, жуте линије на слици лево, прилично успешно пролази кроз дате тачке вероватноћа слова.

У околини очекивања, апсцисе $\mu \approx 8$, нагиб експоненцијалне функције приближно је

$$\frac{d}{dx} f_{\text{par}}(\mu) = -\gamma^2 e^{-\gamma\mu} = -0.006.$$

У поређењу са нагибом праве $\alpha_1 = -0.003$, то је битно стрмије опадање, на уштрб још више занемарених слова ове расподел. Због бржег падања вероватноћа, информација регресије експоненцијалном функцијом је мања од овде претходно обрађених. И заиста, то се види и на попису Шенонових информација:

$$S_y = 4.292664, \quad S_{\text{lin}} = 4.486895,$$
$$S_{\text{par}} = 4.295455, \quad S_e = 4.033481.$$

Због истог очекивања и истих вредности апсцисе, регресије и дате тачке имају исту дисперзију, $\sigma = 11.480287$, али немају исте дистанце $d_{\text{lin}} = 0.072938$, $d_{\text{par}} = 0.037036$ и $d_{\text{e}} = 0.029945$, из чега следи да је експоненцијална расподела најближа датој, мало боља од наведених.

***

Пошто се експоненцијална регресија дистрибуције слова текста показала најбољом, занимљива нам постаје теорема из наслова „41. Ограничења“, да ће са унапред датим очекивањем највећу информацију имати густина вероватноће експоненцијалне расподеле. Расподела слова текста са смисленим садржајем је експоненцијална, а таква расподела „нема памћење“ [64.1. Теорема 1] и отуда закључак да претходни делови текста не утичу битно на следеће.

За разумевање овог парадокс, понудићу предлог. Процес који је „затворен“ са експоненцијалном расподелом већ је тако „пребукиран“ да не може повећавати своју информацију. Успут не нараста, а како (угавном) стару информацију не губи – он далеко стиже непромењен. Експоненцијални су они процеси, дакле, који информације добро конзервирају.

Сада се необична, да не кажем фантастична, представа претходних теорема почиње слагати попут коцкица мозаика. Надам се да примећујете колико је предложено објашњење сагласно са општом тежњом природе ка мањој информацији, да је оштра мисао, извесност теореме, или организација уопште, колико привлачна својим мањком опција свему, толико и одбојна неизвесности.

65.4. О програмирању

Компјутерски језик „Питон“ (Python 3.9.7) користим као лични калкулатор, а понекад програмирам са разним другим поставкама[19]. Радио сам као пројектант и програмер за банке, а и у настави, али то одавно није моје основно занимање. Сложенија „израчунавања“ волим радити са модулима и функцијама које програмирам сам.

Анализу фреквенције слова и цртање графова раде ми две функције, slovaFr() и slovaGr(). Прва има азбуку великих и малих слова, чита задати текст, броји све карактере текста, броји сва дата слова у тексту, сортира фреквенције слова. Обично другом позивам прву, она преноси параметар и враћа листу слова и одговарајућу листу фреквенција. Параметар је бинарни број, рецимо ххх, значења: 1хх да се штампа „текст“, х1х да се штампа „азбука“ детаљно са великим и малим словима, хх1 да се штампа „топ листа“. Ако се не наведе, тај параметар је „000“.

```
def slovaFr(binarno = '000'): # brojanje datih slova u 'teka.txt'
    '''
    binarno = 'bbb', b in {'0', '1'}:
            '1bb' stampa 'text'
            'b1b' stampa 'azbuku'
            'bb1' stampa 'top listu'
    '''

    # open file in read mode
    teka = 'teka.txt'
    file = open(teka, encoding = 'utf-8', mode = 'r')
    text = file.read()
    file.close()

    if binarno[0] == '1': print(text) # ako je binarno '1bb'
    znaka = len(text)

    #eng1 = ['A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J', 'K',
'L', 'M', 'N', 'O', 'P', 'Q', 'R', 'S', 'T', 'U', 'V', 'W', 'X', 'Y',
'Z']
    #eng2 = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j', 'k',
'l', 'm', 'n', 'o', 'p', 'q', 'r', 's', 't', 'u', 'v', 'w', 'x', 'y',
'z']

    srb1 = ['А', 'Б', 'В', 'Г', 'Д', 'Ђ', 'Е', 'Ж', 'З', 'И', 'Ј',
'К', 'Л', 'Љ', 'М', 'Н', 'Њ', 'О', 'П', 'Р', 'С', 'Т', 'Ћ', 'У', 'Ф',
'Х', 'Ц', 'Ч', 'Џ', 'Ш']
    srb2 = ['а', 'б', 'в', 'г', 'д', 'ђ', 'е', 'ж', 'з', 'и', 'ј',
'к', 'л', 'љ', 'м', 'н', 'њ', 'о', 'п', 'р', 'с', 'т', 'ћ', 'у', 'ф',
'х', 'ц', 'ч', 'џ', 'ш']

    azb4 = [] # verovatnoce
    azb3 = [] # komada slova
```

[19] Одговарам на питања.

```
    azb2 = srb2 # mala slova
    azb1 = srb1 # velika slova

    broj = 0
    koda = len(azb1) # broj s slova azbuke
    for i,x in enumerate(azb1):
                slov1 = azb1[i]
                slov2 = azb2[i]
                br1 = text.count(slov1)
                br2 = text.count(slov2)
                br = br1 + br2
                azb3.append(br)
                if binarno[1] == '1': #stampa azbuku, binarno 'b1b'
                        print(slov1, '+', slov2, '=', br1, '+', br2,
'=', br)
                broj = broj + br
                continue

    print('Ukupno svih znakova =', znaka)
    print('Zbirno datih slova =', broj)
    nn = 1
    while True: # buble sort
            if nn == 0: break
            nn = 0
            k = koda - 1
            for i in range(k):
                j = i + 1
                if azb3[i] < azb3[j]:
                        a = azb1[j]
                        b = azb3[j]
                        azb1[j] = azb1[i]
                        azb3[j] = azb3[i]
                        azb1[i] = a
                        azb3[i] = b
                        nn = nn + 1
    rbr = [] # redni broj u 'top listi'
    for i in range(koda): # stampanje 'top liste'
            rbr.append(i+1)
            vv = 0
            if broj != 0: vv = azb3[i]/broj
            azb4.append(vv)
            if binarno[2] == '1': # binarno 'bb1'
                print("%2d. %s %4d  %1.3f" % (rbr[i], azb1[i],
azb3[i], azb4[i]))
    return(azb1, azb4) # vraca top listu: slova, frekvencije
```

То је прва функција. Рецимо, смештена је у фолдер `'D:\Python'`, у модул `'ras'`, када се може позивати и непосредно наредбама:

```
import sys
sys.path.append('D:\Python')
import ras
x,y = ras.slovaFr('111')
```

Имаћете сва три штампања, а у враћеним листама `x`, `y` биће „топ листа“ редом великих слова са одговарајућим фреквенцијама. Прве две наредбе додају „Питону“ позицију за тражење датотека, трећа импортује мој модул, у којем се налазе ове функције, а четврта позива функцију „словних фреквенција“ са параметром „111“.

Следећа је функција slovaGr():

```
def slovaGr(bina1 = '0000', bina2 = '000'): # regresije i graf
        '''
        binarno = 'bb..b', b in {'0, '1'}
        bina1:
                '1bbb' distribucija datih
                'b1bb' linearna regresija
                'bb1b' kvadratna regresija
                'bbb1' eksponencijalna regresija
        bina2:
                '1bb' stampa 'text'
                'b1b' stampa 'azbuku'
                'bb1' stampa 'top listu'
        '''
        slo, fre = slovaFr(bina2)
        s = len(slo)
        x = []
        for i in range(s): x.append(i+1)
        y = fre
        # racun linearne i parabolne regresije
        B0 = B1 = B2 = A1 = A2 = A3 = A4 = 0
        for i in range(s):
                B0 = B0 + y[i]
                B1 = B1 + x[i]*y[i]
                B2 = B2 + y[i]*x[i]**2
                A1 = A1 + x[i]
                A2 = A2 + x[i]**2
                A3 = A3 + x[i]**3
                A4 = A4 + x[i]**4
        print('A1 =', A1)
        print('A2 =', A2)
        print('A3 =', A3)
        print('A4 =', A4)
        print('B0 =', B0)
```

```
print('B1 =', B1)
print('B2 =', B2)
nazivnik = s*A2 - A1**2
brojnik1 = s*B1 - A1*B0
brojnik2 = A2*B0 - A1*B1
alfa1 = alfa2 = 0
if nazivnik != 0:
        alfa1 = brojnik1/nazivnik
        alfa2 = brojnik2/nazivnik
print('f(x) = a1*x + a2')
print('alfa1 =', alfa1)
print('alfa2 =', alfa2)
f1 = []
for i in range(s):
        f1.append(alfa1*x[i] + alfa2)

 # determinante Kramera
import numpy as np
deter0 = np.array([[A4,A3,A2],[A3,A2,A1],[A2,A1,s]])
deter1 = np.array([[B2,A3,A2],[B1,A2,A1],[B0,A1,s]])
deter2 = np.array([[A4,B2,A2],[A3,B1,A1],[A2,B0,s]])
deter3 = np.array([[A4,A3,B2],[A3,A2,B1],[A2,A1,B0]])
det0 = np.linalg.det(deter0)
det1 = np.linalg.det(deter1)
det2 = np.linalg.det(deter2)
det3 = np.linalg.det(deter3)
#print(det0, det1, det2, det3) # stampanje po potrebi
beta1 = beta2 = beta3 = 0
if det0 != 0:
        beta1 = det1/det0
        beta2 = det2/det0
        beta3 = det3/det0
print('f(x) = b1*x^2 + b2*x + b3')
print('beta1 =', beta1)
print('beta2 =', beta2)
print('beta3 =', beta3)

import matplotlib.pyplot as plt # graf distribucije
if bina1[0] == '1':
        plt.scatter(x,y)

from scipy import stats # graf linearne regresije
if bina1[1] == '1':
        nagib, presek, r, p, std_err = stats.linregress(x, f1)
        def funk(x):
                return nagib * x + presek
        gr1 = list(map(funk, x))
```

```
        '''plt.scatter(x,f1) # crta tacke '''
        plt.plot(x, gr1) # crta liniju

import numpy as np # graf kvadratne regresije
x = np.linspace(0, 30, 100)
x = np.arange(0,30,1)
f2 = beta1*x**2 + beta2*x + beta3
if bina1[2] == '1':
        plt.plot(x,f2) # crta parabolu

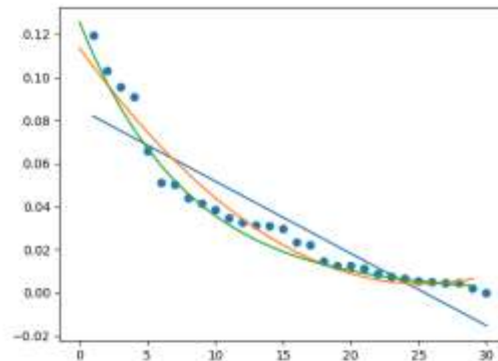
import math # graf eksponencijalne funkcije
gama = 0
if B1 != 0: gama = 1/B1
x = np.linspace(0, 30, 100)
x = np.arange(0,30,1)
f3 = gama*math.e**(-gama*x)
if bina1[3] == '1':
        plt.plot(x,f3) # crta eksponencijalnu

plt.show()
par = [B1, alfa1, alfa2, beta1, beta2, beta3, gama]
return(slo, fre, par) # slova, frekvencije, parametri
```

Овакве програме имам у разним верзијама (у модулима), за различите функције и представљања. Ако покренете наведени графички модул наредбом

```
x, y = ras.slovaGr('1111')
```

дакле са укључена сва четири за цртањем, добићете слику десно, а низови `x` и `y` биће слова и одговарајуће фреквенције у редоследу „топ листе“. Плаве тачке су дате фреквенције, права линија регресије је плава, регресија параболом је жута, а четврта експоненцијална је зелена.



Изостављањем прве јединице (0111) нема тачака, друге (1011) нема праве линије, треће (1101) нема параболе, а четврте (1110) нема експоненцијалне регресије. Покушајте даље комбиновати!

66. Друга азбука

Стил који добар писац има, или тема којој се учењак посветио чине њихов текст читаоцу упућен отприлике константним. Те записане мисли упућене свету, примећене и читане, не могу побећи законитостима канала преноса, верујем, а само је питање да ли је време или наша моћ таква да процесе и формализујемо.

Из претходних разматрања могли смо научити разлике између једноликих, благо косих и веома стрмих графичких расподела вероватноћа слова, а можда такође и речи или идеја. Приметимо ли експоненцијалну зависност тих „случајних“ (на сасвим случајан начин могу се појављивати и неки процеси у посебним ситуацијама, попут децимала броја $\pi = 3.14159$ ... када их први пут читамо) токова, сада знамо да су независни од своје прошлости и зато отпорнији на сметње и промене.

Поред тога, да су „јасније“ мисли мање информативне, да су „расуте“, зато мутне и неразумљиве јер су оптерећене дезинформацијама, да ће препричавањем од уха до уха, кроз генерације такве еволуирати у „својствене“ средини која их преноси чинећи тај дуги ланац „црном кутијом“. Али остаје отворено питање да ли и шта нам облик графа расподеле дубље казује о емисији, односно о језику, аутору, или идејама које се емитују. Може ли тај облик графа бити врста отиска прста, тј. нешто попут ДНК емисије?

Каква је разлика између идеја којима се бавимо обзиром на језик којим говоримо и шта можемо рећи о доменима нашег бављења – писаним језиком ограниченим? Какве би биле разлике између текстуалног, сликовног, музичког изражавања, у расподелама симбола њихових „језика“? Назиру се многа слична питања која ми тему на којој сам се затекао чине Пандорином кутијом.

Није ни то крај приче о теорији према којој је информација ткиво простора, времена и материје, а њена суштина неизвесност. Пре свега, јер је у таквој теорији „истина“ увек одраз „стварности“, па су онда и саме бесконачности (због своје математичке логичности) део неке „реалности“, додуше нама апстрактне, али дохватљиве нашем разуму и неизбежне у баратању са уобичајеним питким коначностима. Не помишљам у тим „фантазијама“ заобићи нити проблеме које су нам у схватању света направили Раселов парадокс (нема скупа свих скупова), Геделова теорема немогућности, или поменути „принцип неизвесности“ примењен на „целу“ васиону.

Другачији, а једнако „чудан пример“ је начин ширења свемира условљен преносом информација из његове прошлости у садашњост и будућност. Сметње које су при томе нужне, очито је да споро настају, јер галаксије далеке прошлости још увек „јасно“ видимо. Међутим, извесност адаптације те информације о стању некадашњости још даљу прошлост учиниће нам „невидљивом“. Не само да ће „Велики прасак“ (пре око 13.8 милијарди година) бити у непрозирној магли за фотоне какве их познајемо, већ да би и физика још даље прошлости могла бити нама невидљива, несхватљива и, штавише, здравом разуму неприхватљива.

Испоставиће се да ова „кратка скрипта“, како ми је изгледала на почетку писања „Информатичке теорије“ , не само да се не може завршити на првом делу, него да неће стати ни са другим. А тек сам се припремио расписати се о сепарабилним процесима, па мало озбиљније о континуалним, затим Колмогоровљевим, посебно још о дискретним изворима, асимптотским поделама, а да и не помињем колико тога имам рећи о кодирању. Међутим, све такве теме мораће на чекање, јер преносници порука на којима овде радимо показаће се кореном многих сличних наставака.

66.1. Енглески језик

Размотримо опет исти мој текст „1.4 Феминизација“ из скрипте „Приче о информацији“ којим смо се бавили у наслову „65. Регресија“, али сада преведен на енглески „1.4 Feminization“ у упоредној скрипти „Information Stories“. Бројимо све знакове енглеског текста (4525) и посебно одвајамо сва велика и мала слова енглеског алфабета (3661). Резултат је:

A + a = 0 + 256 = 256
B + b = 2 + 39 = 41
C + c = 3 + 94 = 97
D + d = 1 + 119 = 120
E + e = 2 + 463 = 465
F + f = 4 + 87 = 91
G + g = 1 + 58 = 59
H + h = 0 + 175 = 175
I + i = 7 + 312 = 319
J + j = 0 + 2 = 2
K + k = 0 + 19 = 19
L + l = 3 + 112 = 115
M + m = 1 + 96 = 97
N + n = 3 + 302 = 305
O + o = 1 + 294 = 295
P + p = 1 + 64 = 65
Q + q = 0 + 6 = 6
R + r = 0 + 230 = 230
S + s = 1 + 243 = 244
T + t = 10 + 354 = 364
U + u = 1 + 99 = 100
V + v = 0 + 48 = 48
W + w = 4 + 65 = 69
X + x = 0 + 10 = 10
Y + y = 0 + 54 = 54
Z + z = 0 + 15 = 15

Након сортирања по броју појављивања слова, добијамо њихов „пласман“:

1. E 465 0.127
2. T 364 0.099
3. I 319 0.087
4. N 305 0.083
5. O 295 0.081
6. A 256 0.070
7. S 244 0.067
8. R 230 0.063
9. H 175 0.048
10. D 120 0.033
11. L 115 0.031
12. U 100 0.027
13. C 97 0.026
14. M 97 0.026
15. F 91 0.025
16. W 69 0.019
17. P 65 0.018
18. G 59 0.016
19. Y 54 0.015
20. V 48 0.013
21. B 41 0.011
22. K 19 0.005
23. Z 15 0.004
24. X 10 0.003
25. Q 6 0.002
26. J 2 0.001

У последњој колони су фреквенције, тј. „статистичке вероватноће“. То су бројеви претходне колоне дељени константним бројем свих слова (3661). Означимо као и пре редне бројеве ове листе са $\vec{x} = (1,2,\dots,26)$, одговарајуће фреквенције $\vec{y} = (0.127, 0.099, \dots, 0.001)$, а са $s = 26$ број слова овог алфабета.

Коефицијенти потребни за израчунавање регресија су $B_0 = \sum_{k=1}^{s} y_k = 1$ и даље редом:

$$B_1 = \sum_{k=1}^{s} x_k y_k = 7.231904 = \mu$$

$$B_2 = \sum_{k-1}^{s} x_k^2 y_k = 82.939634$$

$$A_1 = \sum_{k=1}^{s} x_k = 351$$

$$A_2 = \sum_{k=1}^{s} x_k^2 = 6201$$

$$A_3 = \sum_{k=1}^{s} x_k^3 = 123201$$

$$A_4 = \sum_{k=1}^{s} x_k^4 = 2610621$$

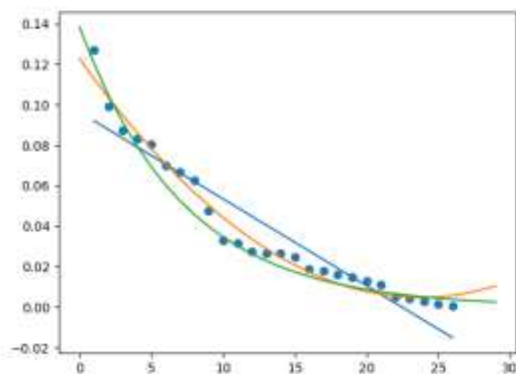
На претходно дате начине, помоћу ових коефицијената добијамо једначине линеарне, квадратне и експоненцијалне регресије датих фреквенција:

$$f_{\text{lin}} = \alpha_1 x + \alpha_2 \qquad \alpha_1 = -0.004286, \quad \alpha_2 = 0.096321$$

$$f_{\text{par}} = \beta_1 x^2 + \beta_2 x + \beta_3 \qquad \beta_1 = 0.000209, \quad \beta_2 = -0.009923, \quad \beta_3 = 0.122625$$

$$f_{\text{e}} = \gamma e^{-\gamma x} \qquad \gamma = \frac{1}{B_1} = 0.138276$$

Графови регресија су:



Плаве су тачке дате расподеле слова енглеског текста, плава линија им је линеарна регресија, жута је параболна, а зелена експоненцијална.

На први поглед нису видљиве иоле битније разлике од одговарајућих прошлих расподела „истог“ српског текста. Поред овога:

$$S_y = 4.134927, \quad S_{\text{lin}} = 4.274688,$$
$$S_{\text{par}} = 4.146303, \quad S_e = 3.856565.$$

Види се да је експоненцијална расподела опет најмање информативна, а линеарна највише. Мањи бројеви информација у односу на оне српског језика пре свега су последица мањег броја слова. Исте мање вредности информације добијале би се и за потпуно насумичне текстове, наиме:

$$S_s = -\sum_{k=1}^{s} p_k \log_2 p_k = -\sum_{k=1}^{s} \frac{1}{s} \log_2 \frac{1}{s} = \log_2 s$$

када су све вероватноће једнаке $p_k = 1/s$, а има их s. Међутим, $\log_2 26 < \log_2 30$, па је зато

$$S_{26} < S_{20}.$$

Подсећам да сам због овога, уопште, наглашавао да природа поред „избегавања“ равноправности више „воли“ и краће расподеле вероватноћа (мањег броја s), јер „тежи“ мањој информацији, када год то може. Са друге стране стоје принципи многострукости, свеприсутности, као и конзервације (закона одржања) информације, па овај свет не нестаје.

Видели смо да су све четири средње вредности (математичка очекивања) оваквих расподела исте (овде $\mu = B_1 = 7.231904$), па су им онда једнаке и варијансе $\sigma^2 = 95.539029$, а зато и дисперзије $\sigma = 9.7744068$. Оне су овде мере средње удаљености координата апсцисе (x-осе) од исте фиксне тачке (очекивања μ) у свим примерима регресија. Међутим, удаљености дате расподеле:

$$d_{\text{l}} = 0.060087, \quad d_{\text{p}} = 0.027475, \quad d_{\text{e}} = 0.033720.$$

нису једнаке. Прве две су мање је од одговарајућих српског језика, али је трећа лошија.

61.2. Лорем ипсум

Лорем ипсум је текст коју чува места обично у графичкој, штампаној и издавачкој индустрији за преглед распореда и визуелних макета. То је лажни текст који се поставља на празна места током припреме штампаног, графичког или веб дизајна. Одломак се приписује непознатом слагачу из 15. века за кога се сматра да је шифровао делове Цицероновог „De Finibus Bonorum et Malorum“, ради употребе у књизи примерака типова. Обично почиње са:

“Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua...”

Лорем ипсум је намерно дизајниран да нема значење, али да изгледа као прави текст, што га чини савршеним чуваром „места“. Погледајмо даље шта може казати статистика о таквом. Десет пасуса текста „лорем ипсума“, које сам насумично генерисао за ову прилику, садржи укупно 7195 знака и од тога 5981 слова енглеског алфабета ($s = 26$), без слова K, W, X, Y, Z:

1.	E	716	0.120	10.	M	281	0.047	19.	F	49	0.008
2.	I	616	0.103	11.	C	258	0.043	20.	H	19	0.003
3.	S	553	0.092	12.	O	218	0.036	21.	J	1	0.000
4.	U	536	0.090	13.	D	198	0.033	22.	K	0	0.000
5.	T	525	0.088	14.	P	143	0.024	23.	W	0	0.000
6.	A	484	0.081	15.	V	107	0.018	24.	X	0	0.000
7.	N	381	0.064	16.	G	86	0.014	25.	Y	0	0.000
8.	L	344	0.058	17.	Q	83	0.014	26.	Z	0	0.000
9.	R	325	0.054	18.	B	58	0.010				

Када се та слова (K, W, X, Y, Z) избаце, добија се скраћена топ листа ($s = 21$):

1.	E	716	0.120	8.	L	344	0.058	15.	V	107	0.018
2.	I	616	0.103	9.	R	325	0.054	16.	G	86	0.014
3.	S	553	0.092	10.	M	281	0.047	17.	Q	83	0.014
4.	U	536	0.090	11.	C	258	0.043	18.	B	58	0.010
5.	T	525	0.088	12.	O	218	0.036	19.	F	49	0.008
6.	A	484	0.081	13.	D	198	0.033	20.	H	19	0.003
7.	N	381	0.064	14.	P	143	0.024	21.	J	1	0.000

За линеарну регресију $f_l(x) = \alpha_1 x + \alpha_2$ налазимо $\alpha_1 = -0.005767$ и $\alpha_2 = 0.1110556$. Квадратна регресија, парабола $f_p(x) = \beta_1 x^2 + \beta_2 x + \beta_3$ је са параметрима $\beta_1 = 0.000167$, $\beta_2 = -0.009438$, $\beta_3 = 0.125126$. Средње вредности (очекивања) једнаке су им $\mu = 6.559438$, а отуда коефицијент $\gamma = 1/\mu = 0.152452$ експоненцијалне регресије $f_e(x) = \gamma e^{-\gamma x}$.

1.	0.120	0.105	0.116	0.131	8.	0.058	0.065	0.060	0.045	15.	0.018	0.025	0.021	0.015
2.	0.103	0.100	0.107	0.112	9.	0.054	0.059	0.054	0.039	16.	0.014	0.019	0.017	0.013
3.	0.092	0.094	0.098	0.096	10.	0.047	0.053	0.047	0.033	17.	0.014	0.013	0.013	0.011
4.	0.090	0.088	0.090	0.083	11.	0.043	0.048	0.042	0.028	18.	0.010	0.007	0.009	0.010
5.	0.088	0.082	0.082	0.071	12.	0.036	0.042	0.036	0.024	19.	0.008	0.001	0.006	0.008
6.	0.081	0.076	0.075	0.061	13.	0.033	0.036	0.031	0.021	20.	0.003	-0.004	0.003	0.007
7.	0.064	0.071	0.067	0.052	14.	0.024	0.030	0.026	0.018	21.	0.000	-0.010	0.001	0.006

На табели је топ листа са упоредним вероватноћама дате, линеарне, квадратне и експоненцијалне расподеле редом. Због једнаких очекивања и једнако баждарених апсциса, све четири расподеле имају исте варијансе $\sigma^2 = 56.385256$ такође, а онда једнаке и дисперзије $\sigma = 7.509012$.

Информације ових расподела:

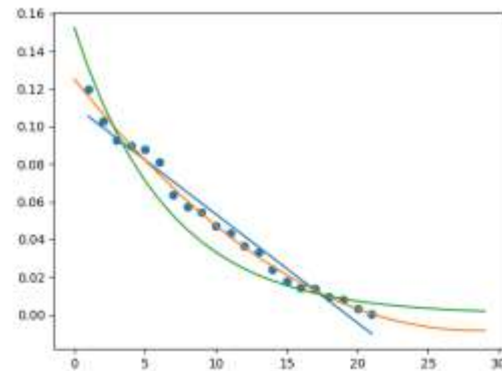
$$S_y = 4.316589, \quad S_l = 3.987831,$$
$$S_p = 3.952735, \quad S_e = 3.617013,$$

показују уједначеност и, очекивано, највећу у случају самих расутих датих тачака и најмању код нај стрмије (у почетку) експоненцијалне расподеле. Међутим, максимална могућа информација ових слова ($s = 21$) је $S_0 = \log_2 21 = 4.392317$, што значи да до S_y једва да има „места“.

Та мала разлика информација, максимална и слова лорем ипсума, $\Delta S = S_0 - S_y = 0.075728$, је „простор“ у којем се ипак могу наћи неке поруке. На пример, могуће је да на мом сајту остављам наводно лажни текст лорем ипсума, а да заправо искористим ту мрву његове не-насумичности за неко кодирање, које би обучен читаоц негде у свету знао декодирати. Да сви остали имају отворен пристум нашим страницама, у уверењу да су бесмислене, док нас два тако комуницирамо.

Удаљености израчунатих расподела од дате:

$$\begin{cases} d_l = 0.028599, \\ d_p = 0.013908, \\ d_e = 0.047638, \end{cases}$$



показују да је за слова лорем ипсума најбоља параболна (квадратна) апроксимација, а да је најлошија експоненцијална. То се помало и види на упоредном графу ових дистрибуција, десно. Опет су плаве тачке дате расподеле, а плава је праволинијска регресија, жута је квадратна. Зелена апроксимација је експоненцијална и најлошија је од ових за лорем ипсум.

Иначе, због веома несразмерних јединица дужина апсцисе и ординате (x и y осе) чини се да је линија ових тачака много више нагета него што јесте. Коефицијент праве је -0.006 само, а то је мање од 1 одсто!

61.3. Вижнерова шифра

Вижнерово шифровање је метода замене азбучног текста коришћењем серије Цезарових шифара, заснованих на словима кључа. Ово је прости облик шифре полиалфабетске замене. На тај начин је кодиран (Vigenere) претходни енглески текст [66.1] са кључном речи „ponedeljak“, који је превод српског текста [65.]. Намерно узимам три верзије истог текста, овде детаљно анализиране, ради лакшег разумевања даљих резултата.

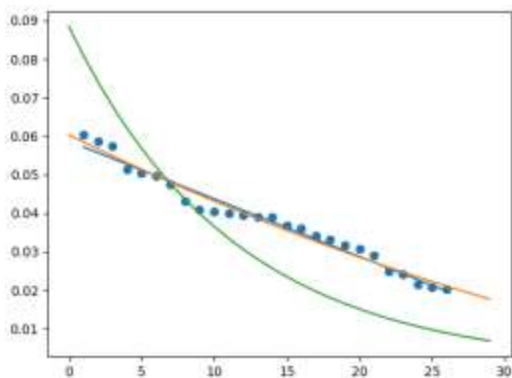
Свих знакова тог текста је 4518, а кодних слова 3661, а за разлику од некодираног енглеског без малих слова, као што се види из пописа резултата у следећој несортираној табели. Слова збирно нађених тамо и овамо има тачно једнако, али нешто од осталих знакова уникод не препознаје.

A + a = 125 + 0 = 125
B + b = 132 + 0 = 132
C + c = 184 + 0 = 184
D + d = 143 + 0 = 143
E + e = 182 + 0 = 182
F + f = 88 + 0 = 88
G + g = 135 + 0 = 135
H + h = 158 + 0 = 158
I + i = 188 + 0 = 188
J + j = 74 + 0 = 74
K + k = 79 + 0 = 79
L + l = 144 + 0 = 144
M + m = 116 + 0 = 116
N + n = 121 + 0 = 121
O + o = 150 + 0 = 150
P + p = 143 + 0 = 143
Q + q = 112 + 0 = 112
R + r = 221 + 0 = 221
S + s = 210 + 0 = 210
T + t = 146 + 0 = 146
U + u = 91 + 0 = 91
V + v = 148 + 0 = 148
W + w = 215 + 0 = 215
X + x = 174 + 0 = 174
Y + y = 106 + 0 = 106
Z + z = 76 + 0 = 76

Већ из ове табеле видимо уједначеност свих могућих ($s = 26$) слова, што је нетипично за текстове са смислом. Ова уједначеност се још боље види након сортирања:

1.	R	221	0.060
2.	W	215	0.059
3.	S	210	0.057
4.	I	188	0.051
5.	C	184	0.050
6.	E	182	0.050
7.	X	174	0.048
8.	H	158	0.043
9.	O	150	0.041
10.	V	148	0.040
11.	T	146	0.040
12.	L	144	0.039
13.	D	143	0.039
14.	P	143	0.039
15.	G	135	0.037
16.	B	132	0.036
17.	A	125	0.034
18.	N	121	0.033
19.	M	116	0.032
20.	Q	112	0.031
21.	Y	106	0.029
22.	U	91	0.025
23.	F	88	0.024
24.	K	79	0.022
25.	Z	76	0.021
26.	J	74	0.020

Од највеће (R: 0,060) до најмање (J: 0,020) фреквенције, тј. „статистичке вероватноће“, необично је мала разлика. За линеарну регресију



$$f_l(x) = \alpha_1 x + \alpha_2$$

два су параметра:

$$\alpha_1 = -0.001497, \quad \alpha_2 = 0.058672.$$

За квадратну регресију

$$f_p(x) = \beta_1 x^2 + \beta_2 x + \beta_3$$

параметри су:

$$\beta_1 = 0.000013,$$
$$\beta_2 = -0.001849,$$
$$\beta_3 = 0.060315.$$

Средња вредност (математичко очекивање) једнако је свим овим расподелама

$$\mu = 11.310571,$$

а отуда коефицијент $\gamma = 1/\mu = 0.088413$ експоненцијалне регресије $f_e(x) = \gamma e^{-\gamma x}$. На тој слици лево, приказани су графови. Експоненцијалну расподелу (зелена линија) видимо као промашај.

Информације ове четири расподеле редом су:

$$S_y = 4.636296, \quad S_l = 4.637270,$$

$$S_p = 4.638081, \quad S_e = 3.989773.$$

Информација потпуно насумичних ($s = 26$) слова била би $S_0 = \log_2 26 = 4.700440$ и то је највише што та слова могу остварити. Дакле, постоји разлика $\Delta S = S_0 - S_y = 0.0638104$ која дозвољава декодирање Вижнерове шифре. Та је разлика мања од сличне лорем ипсума (0.075728), а што год је она мања то је порука дубље сакривена. Са малим узорком текста Вижнеровог кода намогуће је декодирање, он је тада „непробојна шифра“ (назив који је тај код дуго носио). Међутим, са дужим тако кодираним текстовима има начина дешифровања и они су познати.

Дистанце ове три расподеле од дате су:

$$d_l = 0.000099, \quad d_p = 0.000088, \quad d_e = 0.003573.$$

Праволинијска расподела (d_l) најближа је датој, а експоненцијална (d_e) је најдаља. Томе треба додати да је коефицијент опадања ове праве једва један промил, $\alpha_1 = -0.001497$, а то је тако мали број (угао) да би у координатном систему једнаких јединица дужина (апсице и ординате) једва приметили њену нагнутост.

На крају приметимо да су математичко очекивање (средња вредност) и дисперзија (средње одступање од средње вредности) дате расподеле:

$$\mu = 11.310571, \quad \sigma = 7.813040.$$

Средња вредност распона $s = 26$ једнако вероватних слова била би $\mu_0 = s/26 = 13$, а средње одступање од такве средње вредности, на начин ове дисперзије израчунато, било би:

$$\sigma_0 = \sqrt{\frac{(1-\mu_0)^2 + (2-\mu_0)^2 + \cdots + (s-\mu_0)^2}{s}} = 7.516648,$$

што су незнатне разлике од обрађене расподеле. То такође потврђује тежину њеног декодирања, да је понашање Вижнерове шифре блиско „црној кутији“.

67. Шредингерова мачка

Прича о Шредингеровој мачки надрасла је своју првобитну мисију, као и сваки занимљив пример заснован на тачној науци. Смишљена 1935. од стране аустријског физичара Ервина Шредингера да би као мисаони експеримент осветлила тешкоће тумачења тада нове квантне теорије, она је овде згодан додатак сагледавања уопште (апстрактних, математичких) канала преноса информација.

Квантна физика каже да објекат њеног проучавања, попут честице или атома, нема стварност која се може утврдити док се не измери. До мерења, својства предмета, као што је импулс, кодирана су математичком формулом таласне функције о могућим исходима након мерења. Ово кодирање, или (познаваоцима) расветљавање, међутим, наглашава нејасност тих стања пре мерења.

Од одговора понуђених 1930-их, најдаље је добацила тзв. Копенхагенска интерпретација, коју су у данском граду развили светитељи квантне теорије, Нилс Бор и Вернер Хеисенберг. Утврдили су да заиста не постоји дефинитивна реалност пре мерења, када је објекат био у недефинисаном стању познатом као суперпозиција. Штавише, Хајзенбергов приницп неодређености, који је откривен тих година, каже да не можемо сазнавати инхерентне природне параметре честице, посебно положај и импулс, или тренутак и енергију, истовремено и тачно. Каже да то није због слабе технике, малог знања, или сличних наших ограничења, него због дубљих природних разлога.

Шредингеров мисаони експеримент истражује како се све ово дешава када се квантни објекат пренесе на нешто познатије. Замислио је кутију у којој се налази радиоактивни атом, упакована отровна супстанца и мачка. Радиоактивност атома може или не у било ком тренутку учинити да отров убије мачку. Ако је копенхагенско тумачење тачно, онда су и пре него што је дошло до било каквог мерења, атоми, а тиме и мачка, били у суперпозицији жива-или-мртва мачка. Отварањем кутије дешава се мерење и дефинитивно изјашњавање да је мачка или жива или је мртва.

Развој догађаја у кутији, отварање кутије и стање мачке након процес је и има смисла применити на њега сазнања из ове скрипте, ма како нам се то на први поглед чинило апсурдним. На пример, како на микро нивоу постоје објективне неизвесности, које се у неким тачкама прекретнице могу драстично одражавати на макро плану. Или да будућност може утицати на прошлост[20].

Ако заиста постоје неизвесности и дубљих разлога од оних које имамо због својих ограничења, то повлачи да је периодичности [59. Периодични сигнали] требамо сматрати ограниченим појавама. Сви реални процеси тада временом постају „црне кутије“ [61.2. Ергодичка теорема] за евентуалне далеке „читаоце“ прошлости. Такође, да можемо прихватити неку врсту реалности опција које се нису десиле (Dimensions).

Ако смо данас живи, то се стање тада може посматрати и као основа за искључивање мноштава могућности из претходне прошлости, каква је, рецимо, да смо јуче настрадали у саобраћају. Оно ће се показати једнако „објективним“, тј. ослонцем у научним доказима као, на пример, данашње стање смрти некога ко јуче јесте настрадао у саобраћају. Исто би као врсту истине морала видети и сама математика, било у данас познатим а несхваћеним моделима, или у још увек неоткривеним. Онда је важно и питање реалности обрнутог тока времена, макар то било само за врсте „честица“ нама данас, или заувек, потпуно непознате.

[20] Следе не само моја запажања оваквих апсурда.

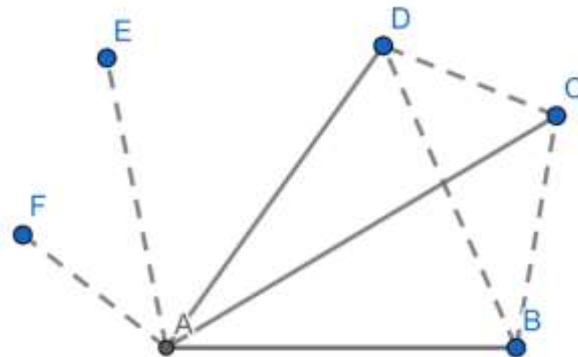
67.1. Дирихлеов принцип

У својој првобитној употреби Дирихлеов принцип[21] каже да ће се у бар једној кућици за голубове наћи бар два голуба, ако има више голубова него кућица. Ако је $n + 1$ предмета распоређено у n посуда ($n = 1,2,3, \dots$), тада ће бар једна посуда садржавати бар 2 предмета, додајемо. У скупу са три природна броја, бар два броја су исте парности (оба парна, или оба непарна). У свакој групи од 13 особа, бар две особе имају рођендан у истом месецу.

Иза ових лакших примера, наизглед безазленог Дирихлеовог принципа, можемо наслутити много теже задатке и проблеме, због чега је он честа тема математичких такмичења на највишем нивоу. На пример, ако је у одељењу 27 ученика и 13 рачунара, онда постоји бар један рачунар за којим су бар три ученика.

Пример 1. Шест градова су повезани сваки са сваким једном директном линијом, аутобусом или железницом, али не обема. Докажимо да међу њима постоје три града која су међусобно везана линијама исте врсте.

Решење. Посматрајмо слику десно. Нека су A, B, C, D, E, F поменутих шест градова. Нека су путеви AB, AC, AD, AE, AF, од којих 5 су бар три исто повезани (аутобусом или железницом). Нека су то AB, AC и AD. Везани су рецимо аутобуском линијом. Ако и BC има аутобуску линију, онда је задатак решен (AB, BC и CA имају аутобуску линију). А ако не, онда је BC железничка веза.



Међутим, такво закључивање важи за линије CD и DB, које нису аутобуске – ако са A нису са линијом исте врсте, па су онда те железничке и исте су врсте везе са градом B. ❑

Нећемо се даље бавити самим Дирихлеовим принципом, него његовим оспоравањем. Наиме, по копенхагенској интерпретацији квантне механике, како је на претходној страни наведено, квантне честице пре мерења биће у стању суперпозиције, неке врсте лимбо[22] стварности, као Чардак ни на небу ни на земљи[23].

Модел сличног стања у „реалности математике“ имамо у тумачењима бесконачности. Издвојићу Хилбертов парадокс хотела Гранд (1924), мисаони експеримент који илуструје „неприхватљиво“ својство бесконачних скупова. Замислимо хипотетички хотел са пребројиво бесконачним бројем соба, од којих су све заузете. Замислимо да тада дође још један гост.

Управник хотела ће и новог госта сместити у собу, а да сви остали имају смештај. Он ће новог слати у собу 1, а госту из собе 1 да оде у собу 2, оном из 2 да оде у 2 и тако даље. Сви гости имаће собе. А дође ли n нових гостију, он ће их упутити у првих n соба, претходнике у следећих n, од $n + 1$ до собе $2n$, ове у следећих n соба и тако даље.

---

[21] Gustav Lejeune Dirichlet (1805 – 1859) немачки математичар.
[22] Лимбо – изван стварности, или хришћанско место између неба и пакла, .
[23] Бајка Вука Стефановића Карџића (1853).

Такође је могуће примити још пребројиво бесконачно много гостију, тако што старог госта из собе 1 упутимо у собу 2 и уопште сваког из n-те собе у собу број $2n$, тако да собе непарног броја остају празне за примање бесконачног низа нових гостију.

Модел је модел, рећи ћете за овакве математичке заврзламе, а стварност је стварност. Међутим, у теорији информације коју заступам, истина је стварност, па је и бесконачност предмет интеракција са нечим „стварним“, без обзира да ли и ми са таквим имали непосредну интеракцију. То ће бити прича о развоју науке, верујем, каква је била о откривању електромагнетног зрачења, од видљиве нам светлости до спектра већином нама невидљивог. Такође о егзистенцији звука, од фреквенција које можемо чути па до схватања звука као треперења средине фреквенцијама које не чујемо.

Трећи пример на исту тему „гурања под тепих“ је „заобилажење“ (Bypass). То је концепт, за сада, само „теорије информације“. Постулирањем егзистенције објективне случајности и подешавањем индуктивне дефиниције дименизије из топологије (Dimension), показује се да трима просторним димензијама треба додати три временске! Идеје долазе почетком 80-их када сам их сматрао само апстрактним „игрицама“ са геометријом, да бих тек неколико деценија касније те „светове“ изван позатог и признатог 4-дим просторно временског континуума, називао псеудо-реалношћу, потом паралелним световима, да бих касније чуо и за „много светова“ квантне механике (Everett, 1957). Разрађивањем је та теорија за мене постајала све прихватљивија.

Са једне стране, полазећи од „математизоване“ идеје света, да оно што можемо доказати да се не може десити – неће се десити, и даље да је истина реалност, а са друге стране, да је информација ткиво реалности, на нашу реалност можемо гледати као на врсту „псеудо реалности“. У пакету са тиме иде и описивање неке „стварности“ комплексним бројевима, који се у разним прорачунима (рецимо збир коњуговано комплексних је реалан број) појављују као реални резултати.

На тај начин гледајући, ако је коначност крајњи домет моћи наших перцепција (не у математици), то још увек не значи да бесконачности, или заобилажења путем паралелне реалности, не постоје. Тада на Дирихлеов принцип можемо гледати као апроксимацију реалности макро света који нас окружује, као врсту последица закона великих бројева, или још шире као на ону врсту „сила“ које настоје да окупљају.

Садашњост се развија током једне временске линије као информација каналом преноса, аналогно начинима трансформације сигнала у својствени преноснику (и преносника сигналу). Препознајемо у томе и одрицања од слобода живих јединки ради сигурности, удруживања у колективе, настанак већих организаме, али и лењост, инерцију, или закое одржања. Зато је наша стварност побегла из бесконачности, а заправо сместила се у једном њеном делу?

67.2. Путање времена

Трајекторије садашњости цртају закони инерције, тачније принцип најмањег дејства, или спонтана тежња развоја стања у мање информативна, у више вероватна. Обрнуто и једнако тачно, теоријска физика рекла би, трајекторије кретања честица решења су Ојлер-Лагранжовим диференцијалних једначина – које су опет на „принципу најмањег дејства“ физике Простор-време[24] (стр. 54).

[24] Space-Time, 1.2.6 Lagrangian, Theorem 1.2.10, p. 53.

Међутим, протицање времена овде сагледавамо мало даље од званичне физике. Пре свега, то је закривљеност 4-дим простор времена теорије релативности у још две димензије времена (Times in Relativistic Motion, November 16, 2014). Сопствени посматрач је делом у времену невидљивом релативном посматрачу и отуда релативистичка дилатација времена.

Брзина протицања времена је (у мојој теорији информације) мања са мањим обимом реализација исхода, са мањом укупном количином догађаја посматраног окружења. Зато ће атомски, или било који други, часовници посматраног система радити спорије када му комуникације теку спорије. А тада је такав мање информативан и, према начелном минимализму, привлачан. Отуда једначине опште релативности показују гравитациону привлачност места са споријим током времена. Зато су и даље тачке система који ротира привлачне оним ближима центру.

Слично се дешава и приликом једноликог инерцијалног кретања (специјалне релативности), када производ јединица времена и дужина у правцу кретања остаје константан, али за посматрача који види кретање оне се мењају тако да су прве веће а друге мање са већом брзином. Сопствени ход је опет делом у временима невидљивим релативном, али такав не повлачи простор за собом (што се у случају гравитације догађа), а нити делује на простор посматрача, па за честице које би могле опажати описана кретања такви дефицити догађаја не постају битни.

Ефекат привлачности мањка информације налазимо и у привлачности лажи (The Truth), због које се оне лакше шире медијима, јер полуистине и фикцију чини привлачнијима за читање од доказа математичких ставова. Наиме, лаж је разблажена истина! Сличне законитости тичу се многих оних теореме о адаптацији поруке која се преноси каналом (и обрнуто је, али мање присутно), да би се након дугог низа таквих константних трансформација порука јављала као својствени вектор, а сам канал као црна кутија.

Када је информација својствени вектор канала, она се не трансформише текући таквим и постиже свој оптимум (минимум) избегавања мењања, а вероватноће које је изражавају свој оптимум (оне максимум). То посуђујемо из линеарне алгебре и наслућујемо из класичних теорија случајности.

Друга страна ове приче је такође занимљива. Спонтаност привлачења мањка „количине опција“ одговара истој одбијања вишка, као што овде откривано у „сили неизвесности“. Универзалност ове појаве мотивише да је тражимо и тамо где никада не бисмо, рецимо у свим врстама физичких сила. Даље у том смеру је запажање да свака данас препозната честица физике има неки набој, да реагује на неку познату физичку силу. Према томе, додати ћемо, све нама познате физичке силе покрећу се сагласно током времена наше садашњости, доследно начелној штедњи информације какву је ми видимо.

Треба бити егоцентричан па веровати да нема других сила само зато што их не можемо видети, ни другачијих честица којима би време текло супротно нашем, нити могућности да „псеудо реалност“ постоји на начине које не можемо опажати непосредно. У томе ће нас већ сада можда поколебати симетрија неких познатих једначина физике на смер времена, а даље то чинити и слободе бирања било које четири од шест димензија простор-времена (нове теорије информације), стављањем дужине ict за временску кооридинату.